DG AND HDG METHODS FOR CURVED STRUCTURES

by

LI FAN

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2013

MAJOR: MATHEMATICS

Approved by:

_____

Co-Advisor Date

_____

Co-Advisor Date

_____

_____

_____

DEDICATION

To My Family

To My Teachers

ACKNOWLEDGEMENTS

First, I express my deepest gratitude to my advisor, Professor Zhimin Zhang. Without his exceptional guidance, constant inspiration, and endless care, I could never have finished this dissertation.

I am grateful to another advisor, Professor Fatih Celiker for not only teaching me how to conduct research but also preparing me for various challenges I will possibly encounter throughout my academic career. He has had an indelible impact in my life.

I am taking this opportunity to thank Professors Sheng Zhang, Hengguang Li and Wen Li for serving in my committee.

During my graduate study at Wayne State University, the faculty members in Department of Mathematics have helped me so much to grow as both a researcher and a teacher, including Professors Guozhen Lu, Daniel Isakson, Robert Bruner, Choon-Jai Rhee, Tze-Chien Sun, George Yin, Mary Klamo. I thank them and all the unnamed ones.

I am indebted to my parents and for their unconditional and unlimited love, support, and inspiration since I was born.

Finally, but not least, I would like to express my appreciation to the entire Department of Mathematics for their hospitality and services. I enjoyed the warm and friendly atmosphere in the department, and I appreciate the support I have received during my study at Wayne State University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

1.1 Introduction to DG method

The Discontinuous Galerkin (DG) method has attracted much attention in the recent years. It was originally developed in [55] for the steady-state neutron transport equation

$$\sigma u + \nabla \cdot (au) = f,$$

where σ is a real constant, $a(x)$ is piecewise constant. DG methods approximate the solution to partial differential equations in finite dimensional spaces spanned by piecewise polynomial base functions. Approximation polynomial spaces are defined without continuity crossing inter-element interfaces, which is different from the way used in traditional conforming and nonconforming finite element methods.

DG methods have been applied to a variety of problems. We refer to [64, 65, 66, 67, 68, 69, 70, 71], for hyperbolic− [72, 75, 76] for parabolic− and [77, 80, 81, 82, 83, 84, 85, 87, 88] for elliptic−partial differential equations. For a fairly thorough compilation of the history of these methods and their applications see [89].

DG methods have several attractive features. They are high order accurate, highly parallelizable (owing to the discontinuous nature of the approximation), very well suited to handling complicated geometries, and in most cases they enjoy an easier treatment, as compared to the conforming finite element methods, of the boundary conditions. Another key advantage of these methods is their compatibility with the adaptivity strategies. This is mainly due to the lack of continuity requirement among different elements which renders the coding of these methods considerably easier for irregular meshes with hanging nodes

as compared to the continuous version of the finite element methods. Moreover, the degree of approximation can easily be changed from one element to another. Therefore, the DG methods are very well suited for *hp*-adaptivity.

1.2 Introduction to Naghdi Arches

For a generally curved thin elastic arch, the Naghdi type arch model determines the transverse deflection w, the normal fiber rotation θ, and the membrane displacement u, all being single variable functions of the arc-length parameter x of the middle curve, by minimizing the functional

$$\frac{1}{2}\int_0^1 [(\theta' + \kappa[u' - \kappa w])^2 + d^{-2}(u' - \kappa w)^2 + d^{-2}(\theta + w' + \kappa u)^2]dx + \int_0^1 (pu + qw)dx \tag{1.1}$$

in a subspace determined by suitable boundary conditions of $[H^1(0,1)]^3$. Here $H^1(0,1)$ is the L^2-based first order Sobolev space. For the simplicity of our notation we have assumed that the model is non-dimensionalized in a way that all the material properties including the Young's modulus, shear modulus, moment of inertia, and the length of the arch are scaled to be equal to one. All the results in this paper can be easily generalized to the case in which they are non-constant functions. The small parameter $d > 0$ represents the dimensionless thickness of the arch. The function κ is x-dependent, and $\kappa(x)$ is the curvature of the middle curve of the arch at the point of coordinate x. The three terms in the first integral respectively represents bending, membrane, and shear effect. When κ is constantly valued, the arch is circular. A straight beam could be viewed as a special arch with $\kappa \equiv 0$, in which case (1.1) decouples to the Timoshenko beam bending model, governing θ and w, and a membrane

model governing u. The functions p and q are the tangential and transverse resultant loads, respectively. Similarly, a displacement vector of a point of the middle curve is decomposed to its tangent component u and normal component w. In Figure 1 we display some of the characteristics of a typical arch. The parametrization is indicated by the mapping that maps

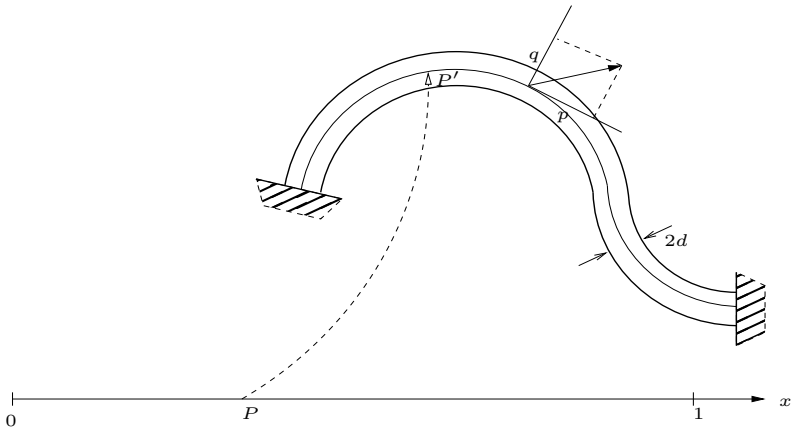

Figure 1: Cross section of an arch clamped at both ends, and arc length parametrization of its middle curve.

$P \in [0, 1]$ to P' on the middle curve. The x coordinate of P is equal to the arc length of the portion of the middle curve from its left end to P'.

For a parameter-dependent model like this, there is the well-known locking issue that indicates the difficulty of accurate computation of the model for small parameter. This problem has been extensively analyzed in the literature, see [11, 12, 35]. Examples of this kind include circular arches [37, 38, 39], the simpler Timoshenko beam bending model [1, 15, 16, 17, 20, 30], and the Reissner–Mindlin plate bending model [3, 4, 6, 8, 9, 10, 13, 21, 25, 26, 27, 29, 31, 32, 33, 40, 41] that does not have the membrane term. This model has as many terms as the Naghdi shell model. In this thesis, we present a DG method for the arch model. It is a extension of the methods for the Timoshenko beam analyzed in [15] and [20].

This is an effort towards the eventual resolution of more challenging shell problems.

Although (1.1) can be used as a starting point to devise DG methods, for the class of the methods we will consider in this thesis, it is more convenient to rewrite it in an equivalent strong form. This does not mean, however, that the variational form (1.1) has lost its significance. Indeed, our proof of existence and uniqueness of the DG approximation as well as its error analysis rely on *energy arguments* inspired by the fact that the DG solution is an approximation to a minimizer of the quadratic functional (1.1). By introducing the scaled membrane stress $N = d^{-2}(u' - \kappa w)$, the scaled transverse shear stress $T = d^{-2}(\theta + w' + \kappa u)$, and the bending moment $M = \theta' + \kappa(u' - \kappa w)$, the Naghdi arch model can be written as a system of first order ordinary differential equations:

$$w' + \theta + \kappa u = d^2 T, \tag{1.2a}$$

$$u' - \kappa w = d^2 N, \tag{1.2b}$$

$$\theta' + \kappa(u' - \kappa w) = M, \tag{1.2c}$$

$$M' = T, \tag{1.2d}$$

$$N' + (\kappa M)' - \kappa T = p, \tag{1.2e}$$

$$T' + \kappa^2 M + \kappa N = q, \tag{1.2f}$$

defined on $\Omega := (0, 1)$. This is a starting point from which one could derive DG methods. In this thesis, we shall derive the DG and HDG methods based on a simplified model that has, as approximations to the elasticity theory, the same accuracy as the Naghdi model. It is proved that if one simplifies the model (1.1) by removing the membrane related term $\kappa[u' - \kappa w]$ from the bending moment M, the model solution will only be changed negligibly. In particular, the solution θ, u, w will deviate in the H^1 norm by $O(d^2)$, and so are the

M, N and T in the L^2 norm. For a detailed explanation of this we refer to [36]. The model thus simplified is often called the mini-model. Consequently, the term $(\kappa M)'$ in (1.2e) and the term $\kappa^2 N$ in (1.2f) can also be neglected without significantly affecting the accuracy of the model. For the sake of brevity and clarity of the presentation and to avoid unnecessary technicalities, we will embrace these simplifications and henceforth work with the following governing equations

$$
\begin{aligned}
w' + \theta + \kappa u &= d^2 T, && \text{(1.3a)}\\
u' - \kappa w &= d^2 N, && \text{(1.3b)}\\
\theta' &= M, && \text{(1.3c)}\\
M' &= T, && \text{(1.3d)}\\
N' - \kappa T &= p, && \text{(1.3e)}\\
T' + \kappa N &= q. && \text{(1.3f)}
\end{aligned}
$$

We note, however, that *all* of the results presented in this paper will remain valid if one chooses to design analogous DG methods based on the original model given by (1.2). To complete the model and ensure the existence and uniqueness of its solution we must impose suitable boundary conditions; we take, for example, the following clamped boundary conditions:

$$
\begin{aligned}
w(0) = w_0, \quad u(0) &= u_0, \quad \theta(0) = \theta_0,\\
w(1) = w_1, \quad u(1) &= u_1, \quad \theta(1) = \theta_1.
\end{aligned}
\tag{1.4}
$$

As will be evident from our analysis, the introduction of the variable curvature function κ and the additional unknowns u and N render the analysis of the numerical methods more challenging. For example, the well posedness of the DG methods requires special conditions

which was not the case for their counterparts for the Timoshenko beam problem. Furthermore, although the error analysis technique is mainly based on the analysis carried out in [20] and [15], the careful reader will notice that there are certain technicalities which do not carry over in a straightforward fashion. Roughly speaking, the main difficulties are caused by the variable nature of the curvature, and the coupling between the transversal (w and T) and tangential (u and N) unknowns. These observations are in agreement with the practical experience that shell structures exhibit more complicated behavior then those of plates.

Finally, we note that although classical continuous Galerkin methods have been developed and analyzed for arch models, it has been shown that [39] in their primal form they suffer from shear and membrane locking. Moreover, to the best of our knowledge, all of the existing methods are limited to circular arches in which case the curvature κ is identical to a constant. It has been shown that [39] the so-called reduced integration technique which is equivalent to certain mixed methods resolves locking. However, the DG framework we introduce and study in this paper offers a more systematic approach and hence is a promising candidate for more challenging problems such as plates and shells. Various advantages of DG methods over other existing methods have been discussed in [2].

The main motivation for considering this simple, one-dimensional model is that it constitutes a stepping stone towards the more challenging goal of devising DG methods for shells. The construction of numerical methods for shells is delicate because, as the thickness of the shell decreases to zero, the numerical method can exhibit what is called in the engineering literature as *shear* and *membrane locking*. Mathematically, this is reflected in the deterioration of the convergence properties of the method as the thickness becomes small. Since some numerical methods for the Naghdi arch model exhibit (shear) locking (as the thickness

of the arch goes to zero), it is instructive to devise locking-free DG methods for this model before considering shells.

Considerable amount of effort has been devoted to the understanding and resolution of shear and membrane locking in structures. Considering the nature of the problem, it is understandable that such effort originated in engineering applications, and was first documented in the engineering literature. The seminal publication in the area, coauthored by Zienkiewicz, Taylor and Too [90], documents the difficulty related to shear effects and uses reduced integration technique to mitigate the problem. The physical understanding of the problem was here critical to devise a remedy, and the resulting technique (reduced integration) is to this day widely used in various commercial software. The term shear locking appears to be coined by Hughes, Taylor and Kanoknukulchai [29] in the context of plate analysis

In parallel with developments related to shear locking, researchers struggled with similar difficulties caused by membrane effects, manifesting themselves in curved structures, such as arches and shells, for example Ashwell and Sabir [73], Lee and Pian [74], Parish [62]. A more thorough explanation of those effects was provided by Belytschko and Stolarski [5], who also introduced the term membrane locking. They subsequently showed that in some models of curved structures there is a delicate interaction between shear and membrane effects, [23].

Over the last two decades or so, there has been a flurry of research activities dealing with shear and membrane locking, and a large number of publications have appeared. Several variations of the known approaches and a number of new ones were developed and described in literature within that time. While related to this work, those approaches address the problem of locking somewhat differently than what we describe here; the interested reader is therefore referred to [78] for a review of many of them. For a locking-free finite element

method for shells we refer to Arnold and Brezzi [7], and for a family of locking-free DG methods for the Reissner-Mindlin plates we refer to Arnold, Brezzi and Marini [4].

While deeply rooted in physical attributes of the analyzed phenomena, locking is essentially a mathematical problem and its challenge was undertaken by mathematicians early on. Arnold [1] proved that shear-locking continuous finite element methods can become locking-free if they are modified by the so-called reduced integration technique. In [30], Li analyzed the p- and hp-versions of the continuous finite element method and proved error estimates independent of the thickness of the beam. These versions of the method take advantage of the extra degrees of freedom gained by increasing the polynomial degree of the approximation. In [37], [38], and [39], Zhang considers circular arch problems. Here shear-locking (and also membrane locking) is again an issue when the arch is thin. Indeed, if the primal form of the method is used where the only unknowns are the displacement and the rotation, both p- and hp- versions exhibit locking. On the other hand, if the shear force is introduced as an additional unknown, along with the membrane forces, and a mixed formulation is employed then both versions can be made free from locking. Following an approach similar to that of Arnold's, Zhang [37, 38, 39] was able to prove error estimates independent of the thickness of the arch.

In [42], the DG methods for the Naghdi arch were introduced and sufficient conditions that ensure the existence and uniqueness of their approximate solutions were proved. Moreover, preliminary numerical experiments were obtained which indicated that, when polynomials of degree p are used, that the optimal order of convergence of $p + 1$ is achieved for the h-version; exponential convergence for the p-version of a DG method was also obtained numerically. Later, in [42], the fact that *all* the numerical traces of the h-version of the DG

method superconverge with order $2\,p+1$ was uncovered and a local post-processing resulting in a uniformly accurate solution of order $2\,p+1$ was devised and numerically tested. These results held uniformly with respect to the thickness of the arch. In this thesis, we put all the above mentioned numerical results on a firm mathematical ground.

2 Locking-free Optimal DG Method for Arches

2.1 The DG Methods for Naghdi Arches

In this section, we introduce the general form of the DG methods. We then provide conditions under which the method is well defined.

2.1.1 The weak formulation for the continuous case

To display the weak formulation we use to define the DG methods, we need to introduce some notation. We begin by partitioning the computational domain into intervals. Given the set of nodes $\mathcal{E}_h := \{x_j\}_{j=0}^{\mathcal{N}}$, where $0 = x_0 < x_1 < \cdots < x_{\mathcal{N}-1} < x_{\mathcal{N}} = 1$, we set $I_j := (x_{j-1}, x_j)$, $h_j := x_j - x_{j-1}$ and $h := \max_{1\le j\le \mathcal{N}}$. We also set $\Omega_h := \cup_{j=1}^{\mathcal{N}} I_j$. Then, we write

$$(f,g)_{\Omega_h} := \sum_{j=1}^{\mathcal{N}} \int_{I_j} fg \quad \text{and} \quad \langle R, [\![f]\!]\rangle_{\mathcal{E}_h} := \sum_{j=0}^{\mathcal{N}} R(x_j)[\![f]\!](x_j).$$

Here, R is any function defined on the set of nodes $\mathcal{E}_h$ and $[\![f]\!]$ is the *jump* of the function f across the nodes which is defined as follows.

$$[\![f]\!](x_j) = \begin{cases} -f(0^+) & \text{for } j = 0, \\ f(x_j^-) - f(x_j^+) & \text{for } 0 < j < \mathcal{N}, \\ f(1^-) & \text{for } j = \mathcal{N}. \end{cases}$$

Here, $f(x_j^{\pm}) := \lim_{\epsilon\downarrow 0} f(x_j \pm \epsilon)$. These jumps are well defined for f in $H^1(\Omega_h)$.

It is now easy to see that if we assume that $(T, N, M, \theta, u, w) \in [H^1(\Omega)]^6$, we have

$$-(w, v_1')_{\Omega_h} + \langle w, [\![v_1]\!]\rangle_{\varepsilon_h} + (\theta, v_1)_{\Omega_h} + (\kappa\, u, v_1)_{\Omega_h} = d^2(T, v_1)_{\Omega_h}, \tag{2.1a}$$

$$-(u, v_2')_{\Omega_h} + \langle u, [\![v_2]\!]\rangle_{\varepsilon_h} - (\kappa\, w, v_2)_{\Omega_h} = d^2(N, v_2)_{\Omega_h}, \tag{2.1b}$$

$$-(\theta, v_3')_{\Omega_h} + \langle \theta, [\![v_3]\!]\rangle_{\varepsilon_h} = (M, v_3)_{\Omega_h}, \tag{2.1c}$$

$$-(M, v_4')_{\Omega_h} + \langle M, [\![v_4]\!]\rangle_{\varepsilon_h} = (T, v_4)_{\Omega_h}, \tag{2.1d}$$

$$-(N, v_5')_{\Omega_h} + \langle N, [\![v_5]\!]\rangle_{\varepsilon_h} - (\kappa\, T, v_5)_{\Omega_h} = (p, v_5)_{\Omega_h}, \tag{2.1e}$$

$$-(T, v_6')_{\Omega_h} + \langle T, [\![v_6]\!]\rangle_{\varepsilon_h} + (\kappa\, N, v_6)_{\Omega_h} = (q, v_6)_{\Omega_h}, \tag{2.1f}$$

for all $v_i \in H^1(\Omega_h)$ for $i = 1, \ldots, 6$. This is the weak formulation we will use to define the DG methods.

2.1.2 The general DG methods

The approximate solution (T_h, N_h, M_h, θ_h, u_h, w_h) given by the DG method is sought in the finite dimensional space $\Pi_{i=1}^6 V_h^{k_i}$ where $V_h^k := \{v : \Omega_h \mapsto \mathbb{R} : \quad v|_{I_j} \in \mathcal{P}^k(I_j), j = 1, \ldots, \mathcal{N}\}$, and $\mathcal{P}^k(K)$ is the set of all polynomials on K of degree not exceeding k. It is determined by requiring that

$$-(w_h, v_1')_{\Omega_h} + \langle \widehat{w}_h, [\![v_1]\!]\rangle_{\varepsilon_h} + (\theta_h, v_1)_{\Omega_h} + (\kappa\, u_h, v_1)_{\Omega_h} = d^2(T_h, v_1)_{\Omega_h} \tag{2.2a}$$

$$-(u_h, v_2')_{\Omega_h} + \langle \widehat{u}_h, [\![v_2]\!]\rangle_{\varepsilon_h} - (\kappa\, w_h, v_2)_{\Omega_h} = d^2(N_h, v_2)_{\Omega_h} \tag{2.2b}$$

$$-(\theta_h, v_3')_{\Omega_h} + \langle \widehat{\theta}_h, [\![v_3]\!]\rangle_{\varepsilon_h} = (M_h, v_3)_{\Omega_h} \tag{2.2c}$$

$$-(M_h, v_4')_{\Omega_h} + \langle \widehat{M}_h, [\![v_4]\!]\rangle_{\varepsilon_h} = (T_h, v_4)_{\Omega_h} \tag{2.2d}$$

$$-(N_h, v_5')_{\Omega_h} + \langle \widehat{N}_h, [\![v_5]\!]\rangle_{\varepsilon_h} - (\kappa\, T_h, v_5)_{\Omega_h} = (p, v_5)_{\Omega_h} \tag{2.2e}$$

$$-(T_h, v_6')_{\Omega_h} + \langle \widehat{T}_h, [\![v_6]\!]\rangle_{\varepsilon_h} + (\kappa\, N_h, v_6)_{\Omega_h} = (q, v_6)_{\Omega_h} \tag{2.2f}$$

hold for all $v_i \in V_h^{k_i}$ for $i = 1, \ldots, 6$.

To complete the definition of the method, we have to define the numerical traces $(\widehat{T}_h, \widehat{N}_h, \widehat{M}_h, \widehat{\theta}_h, \widehat{u}_h, \widehat{w}_h)$ at the nodes. We assume that the general form of these traces is as follows. For an interior node $x_j \in \mathcal{E}_h^\circ := \{x_1, x_2, \ldots, x_{\mathrm{N}-1}\}$, we take

$$\begin{aligned}
\widehat{w}_h &= \{\!\{ w_h \}\!\} + C_{11}[\![w_h]\!] + C_{12}[\![u_h]\!] + C_{13}[\![\theta_h]\!] + C_{14}[\![M_h]\!] + C_{15}[\![N_h]\!] + C_{16}[\![T_h]\!],\\
\widehat{u}_h &= \{\!\{ u_h \}\!\} + C_{21}[\![w_h]\!] + C_{22}[\![u_h]\!] + C_{23}[\![\theta_h]\!] + C_{24}[\![M_h]\!] + C_{25}[\![N_h]\!] + C_{26}[\![T_h]\!],\\
\widehat{\theta}_h &= \{\!\{ \theta_h \}\!\} + C_{31}[\![w_h]\!] + C_{32}[\![u_h]\!] + C_{33}[\![\theta_h]\!] + C_{34}[\![M_h]\!] + C_{35}[\![N_h]\!] + C_{36}[\![T_h]\!],\\
\widehat{M}_h &= \{\!\{ M_h \}\!\} + C_{41}[\![w_h]\!] + C_{42}[\![u_h]\!] + C_{43}[\![\theta_h]\!] + C_{44}[\![M_h]\!] + C_{45}[\![N_h]\!] + C_{46}[\![T_h]\!],\\
\widehat{N}_h &= \{\!\{ N_h \}\!\} + C_{51}[\![w_h]\!] + C_{52}[\![u_h]\!] + C_{53}[\![\theta_h]\!] + C_{54}[\![M_h]\!] + C_{55}[\![N_h]\!] + C_{56}[\![T_h]\!],\\
\widehat{T}_h &= \{\!\{ T_h \}\!\} + C_{61}[\![w_h]\!] + C_{62}[\![u_h]\!] + C_{63}[\![\theta_h]\!] + C_{64}[\![M_h]\!] + C_{65}[\![N_h]\!] + C_{66}[\![T_h]\!],
\end{aligned} \tag{2.3}$$

where $\{\!\{f\}\!\}(x_j) := \frac{1}{2}(f(x_j^-) + f(x_j^+))$. At $x = 0$, we take

$$\begin{aligned}
\widehat{w}_h &= w_0,\\
\widehat{u}_h &= u_0,\\
\widehat{\theta}_h &= \theta_0,\\
\widehat{M}_h &= M_h^+ \quad + C_{41}(w_0 - w_h^+) + C_{42}(u_0 - u_h^+) + C_{43}(\theta_0 - \theta_h^+),\\
\widehat{N}_h &= N_h^+ \quad + C_{51}(w_0 - w_h^+) + C_{52}(u_0 - u_h^+) + C_{53}(\theta_0 - \theta_h^+),\\
\widehat{T}_h &= T_h^+ \quad + C_{61}(w_0 - w_h^+) + C_{62}(u_0 - u_h^+) + C_{63}(\theta_0 - \theta_h^+),
\end{aligned} \tag{2.4}$$

and at $x = 1$,

$$\begin{aligned}
\widehat{w}_h &= w_1, \\
\widehat{u}_h &= u_1, \\
\widehat{\theta}_h &= \theta_1, \\
\widehat{M}_h &= M_h^- \quad + C_{41}(w_h^- - w_1) + C_{42}(u_h^- - u_1) + C_{43}(\theta_h^- - \theta_1), \\
\widehat{N}_h &= N_h^- \quad + C_{51}(w_h^- - w_1) + C_{52}(u_h^- - u_1) + C_{53}(\theta_h^- - \theta_1), \\
\widehat{T}_h &= T_h^- \quad + C_{61}(w_h^- - w_1) + C_{62}(u_h^- - u_1) + C_{63}(\theta_h^- - \theta_1).
\end{aligned} \tag{2.5}$$

This completes the definition of the DG methods.

Note how the boundary conditions are incorporated into the method through the definition of the numerical traces at the border. Note also that the functions C_{ij} defining the numerical traces are not necessarily constant on $\mathcal{E}_h$, and can have different values at different nodes. In the following two subsections, we investigate the role of these functions. In particular, we show that out of these thirty six functions, fifteen can be (and in fact should be) expressed in terms of the remaining twenty one and that only six of them have an impact on the "energy" of the discretization.

2.1.3 The discrete energy identity

To see this, we consider a classical energy argument. It is not difficult to see that if we take

$$v_1 = T, \quad v_2 = N, \quad v_3 = N, \quad v_4 = \theta, \quad v_5 = u, \quad v_6 = w,$$

in the equations (2.1), integrate by parts, and add them, we obtain the energy identity

$$d^2(T,T)_{\Omega_h} + d^2(N,N)_{\Omega_h} + (M,M)_{\Omega_h} \;=\; -(p,u)_{\Omega_h} - (q,w)_{\Omega_h} + bc,$$

where

$$\begin{aligned} bc = \quad & w_1 T(1^-) - w_0 T(0^+) \\ & + u_1 N(1^-) - u_0 N(0^+) \\ & + \theta_1 M(1^-) - \theta_0 M(0^+). \end{aligned}$$

Since this identity captures an essential feature of the problem under consideration, we would like to obtain a similar energy identity for the DG method. Such an identity is obtained in the following result.

Proposition 2.1 (Discrete energy identity)**.** *Assume that* $(T_h, N_h, M_h, \theta_h, u_h, w_h)$ *is a solution of the DG method defined by the weak formulation* (2.2) *and the numerical traces given by* (2.3)*,* (2.4)*, and* (2.5)*. Assume that for* all *nodes* $e \in \mathcal{E}_h$ *we have*

$$\begin{aligned} & C_{66} = -C_{11}, \quad C_{56} = -C_{12}, \quad C_{46} = -C_{13}, \quad C_{36} = -C_{14}, \quad C_{26} = -C_{15}, \\ & C_{65} = -C_{21}, \quad C_{55} = -C_{22}, \quad C_{45} = -C_{23}, \quad C_{35} = -C_{24}, \\ & C_{64} = -C_{31}, \quad C_{54} = -C_{32}, \quad C_{44} = -C_{33}, \\ & C_{63} = -C_{41}, \quad C_{53} = -C_{42}, \\ & C_{62} = -C_{51}. \end{aligned} \tag{2.6}$$

Then, we have

$$\Theta_{interior} + \Theta_{jumps} = \Theta_{loads} + \Theta_{bc}, \tag{2.7}$$

where

$$\Theta_{interior} = d^2(T_h, T_h)_{\Omega_h} + d^2(N_h, N_h)_{\Omega_h} + (M_h, M_h)_{\Omega_h},$$

and

$$\Theta_{loads} = -(p, u_h)_{\Omega_h} - (q, w_h)_{\Omega_h}.$$

Here, setting $C_{16} = C_{25} = C_{34} = 0$ *at the boundary nodes, we have*

$$\Theta_{jumps} = -\sum_{e\in\mathcal{E}_h} \Big(C_{16}[\![T_h]\!]^2 + C_{25}[\![N_h]\!]^2 + C_{34}[\![M_h]\!]^2 + C_{43}[\![\theta_h]\!]^2 + C_{52}[\![u_h]\!]^2 + C_{61}[\![w_h]\!]^2 \Big)(e),$$

and $\Theta_{bc} = \Theta_{bc,1} - \Theta_{bc,0}$, *where*

$$\begin{aligned}
\Theta_{bc,1} = \quad & w_1[T_h(1^-) \ - C_{61}(1)w_h(1^-) - C_{51}(1)u_h(1^-) - C_{41}(1)\theta_h(1^-)] \\
& + u_1\,[N_h(1^-) \ - C_{62}(1)w_h(1^-) - C_{52}(1)u_h(1^-) - C_{42}(1)\theta_h(1^-)] \\
& + \theta_1\,[M_h(1^-) - C_{63}(1)w_h(1^-) - C_{53}(1)u_h(1^-) - C_{43}(1)\theta_h(1^-)],
\end{aligned}$$

and

$$\begin{aligned}
\Theta_{bc,0} = \quad & w_0[T_h(0^+) \ + C_{61}(0)w_h(0^+) + C_{51}(0)u_h(0^+) + C_{41}(0)\theta_h(0^+)] \\
& + u_0\,[N_h(0^+) \ + C_{62}(0)w_h(0^+) + C_{52}(0)u_h(0^+) + C_{42}(0)\theta_h(0^+)] \\
& + \theta_0\,[M_h(0^+) + C_{63}(0)w_h(0^+) + C_{53}(0)u_h(0^+) + C_{43}(0)\theta_h(0^+)],
\end{aligned}$$

Proof. The proof of the above result follows by mimicking what was done for the continuous case, that is, by taking

$$v_1 = T_h, \quad v_2 = N_h, \quad v_3 = M_h, \quad v_4 = \theta_h, \quad v_5 = u_h, \quad v_6 = w_h,$$

in the definition of the DG method (2.2), integrating by parts, adding the resulting equations, and carrying out some algebraic manipulations. □

It is now clear that if we take

$$-C_{16},\ -C_{25},\ -C_{34},\ -C_{43},\ -C_{52},\ -C_{61} \geq 0, \tag{2.8}$$

then each of the terms of Θ_{jumps} can be considered to be an *energy* associated with the discontinuous nature of the discretization. Thus, the above condition ensures that the appearance of the jumps in the DG approximation is accompanied by an increase of the total energy of the system. Since this can also be thought of as being a stabilizing effect, they are called the *stabilization* functions. None of the remaining functions appear in the expression for the energy of the approximation, as we can see in the above result. On the other hand, if we penalize the jumps "too much", then the DG method might behave like a typical continuous method and might lock: it would produce very bad approximations for small values of d. On the contrary, if these penalization parameters are chosen appropriately, the DG method will produce a very good approximation. We illustrate this phenomenon in Figure 2. Therein, we take $p(x) = q(x) \equiv 1$, and $\kappa(x) \equiv 1$, for $x \in \Omega = (0, 1)$, together with homogeneous boundary conditions $w = u = \theta = 0$ on $\partial\Omega = \{0, 1\}$. We also show approximations by two of the DG methods just described. We compute the piecewise linear ($k = 1$) DG approximations for an arch of thickness $d = 10^{-3}$. To better illustrate our point we employ a very coarse uniform mesh of size $h = 0.1$. Both methods take

$$C_{11}(x) = C_{22}(x) = C_{33}(x) = -C_{44}(x) = -C_{55}(x) = -C_{66}(x) = 1/2$$

at all interior nodes $x \in \mathcal{E}_h^\circ$, and all the remaining coefficients equal to zero, except for C_{43}, C_{52}, and C_{61}. The first DG method *strongly* penalizes the jumps of the displacements w and u, and the rotation θ, since it takes

$$C_{43}(x) = C_{52}(x) = C_{61}(x) = -10^6$$

at all nodes $x \in \mathcal{E}_h$.

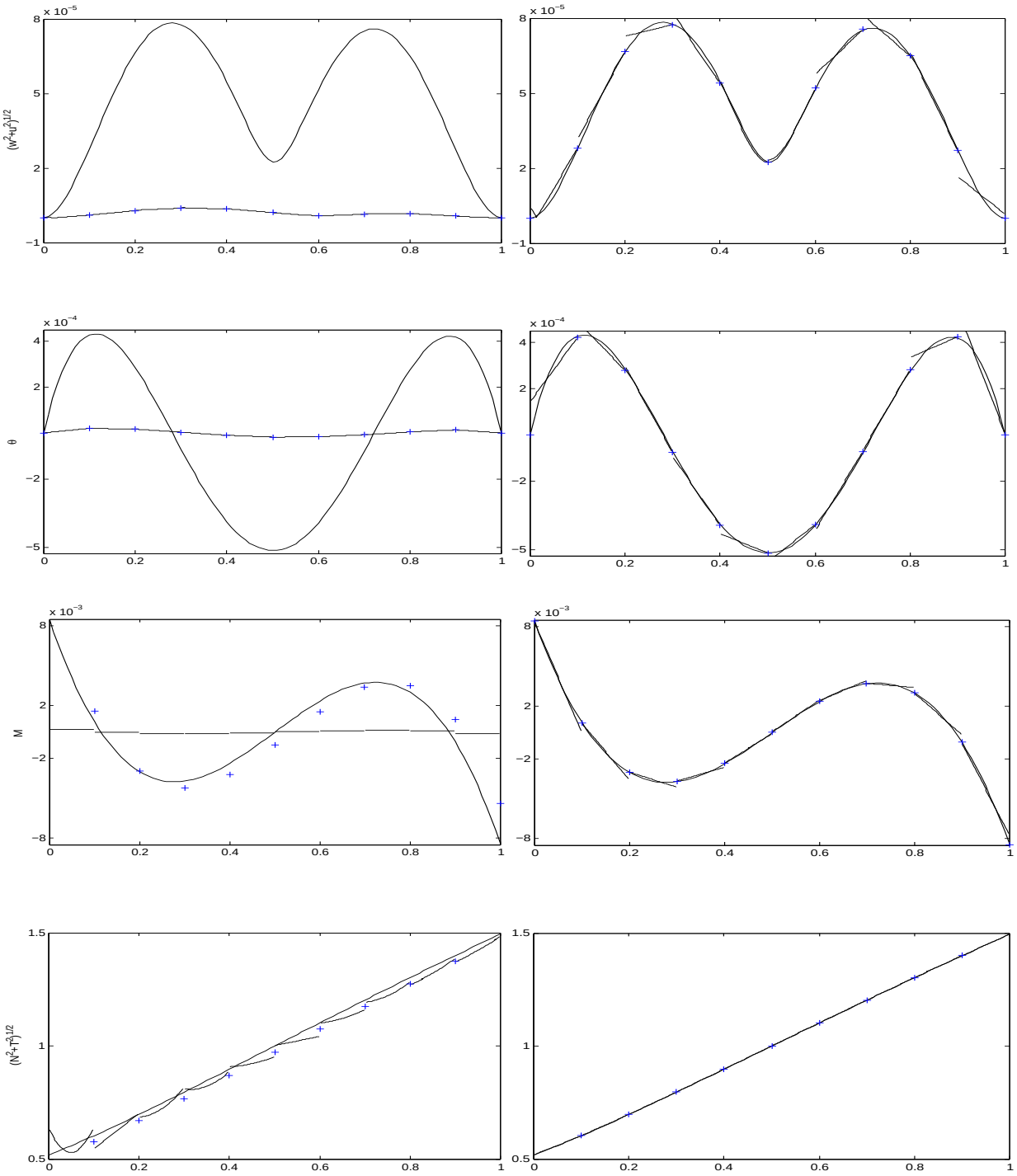

Figure 2: The case $d = 10^{-3}$ and $h = 0.1$. Exact (solid line) and DG approximations (solid line segments and, for the numerical traces, +).

Left column: $C_{43} = C_{52} = C_{61} = -10^6$ at all the nodes. Right column: $C_{43} = C_{52} = C_{61} = 0$ at all the nodes except $C_{43}(1) = C_{52}(1) = C_{61}(1) = -100/h$.

We can see in Figure 2, left column, as expected, it locks. The second method, however, does *not* penalize those jumps *at all* since it takes

$$C_{43}(x) = C_{52}(x) = C_{61}(x) = 0$$

at all nodes except at $x = 1$ where it takes

$$C_{43}(1) = C_{52}(1) = C_{61}(1) = -100/h$$

to *weakly* enforce the boundary conditions there. In Figure 2, right column, we can see that the method produces an excellent approximation of the exact solution. In this paper, we prove that the first method as well converges optimally if the penalization parameters are chosen properly so that the jumps are not superpenalized. We also show that the convergence is independent of the thickness of the arch.

2.1.4 Existence and uniqueness of the DG approximation

The DG method defined by the weak formulation (2.2) and the numerical traces given by (4.2), (2.4), and (2.5) has a unique solution provided that the functions C_{ij}, for $1 \leq i, j \leq 6$, and the polynomial degrees k_i for $1 \leq i \leq 6$, are suitably chosen. The following theorem gives sufficient conditions for this to happen.

Theorem 2.2 (Existence and uniqueness of the DG approximation)**.** *Consider the DG method defined by the weak formulation* (2.2) *and the numerical traces given by* (4.2)*,* (2.4)*, and* (2.5)*. Assume that the conditions* (2.6) *and* (2.8) *are satisfied. Furthermore, suppose that*

$$-C_{43},\ -C_{52},\ -C_{61} > 0 \quad on \quad \mathcal{E}_h, \tag{2.9}$$

and that

$$k_1, k_2 \geq \max\{k_5, k_6\}, \quad k_3 \geq k_4 - 1. \tag{2.10}$$

Then the method has a unique solution provided that

$$h_j \leq \frac{1}{2\|\kappa - \overline{\kappa}_j\|_{L^\infty(I_j)}} \tag{2.11}$$

on the elements I_j where κ is not identically equal to a constant. Here $\overline{\kappa}_j$ denotes the average value of κ on I_j.

A proof of this theorem can be found in Appendix **A**.

2.2 Main Results

For the simplicity of the presentation, in the rest of the paper we restrict ourselves to a particular class of DG methods in which the polynomial degrees k_i are all equal to a given $k \geq 0$ for $i = 1, 2, \ldots, 6$. The functions C_{ij} are defined as follows

$$C_{16} = C_{25} = C_{34} = C_{43} = C_{52} = C_{61} = -\mathsf{c} \tag{2.12}$$

for all x in $\mathcal{E}_h$, except

$$C_{16} = C_{25} = C_{34} = 0 \quad \text{on} \quad \partial\Omega. \tag{2.13}$$

Here, $\mathsf{c} > 0$ is any constant which is independent of the mesh size h. We assume that

$$C_{ij}^2 \leq \mathsf{c} \quad \text{for all } i, j = 1, \ldots, 6, \tag{2.14}$$

and that

$$(C_{ii}(x) - 1/2)^2 \leq \mathsf{c} \quad \text{for all } i = 1, \ldots, 6. \tag{2.15}$$

Such a choice can be obtained, for example, by setting

$$C_{16} = C_{25} = C_{34} = C_{43} = C_{52} = C_{61} = -1$$

for all x in $\mathcal{E}_h$,

$$C_{16} = C_{25} = C_{34} = 0 \quad \text{on} \quad \partial\Omega,$$

and setting all the remaining C_{ij}'s to zero.

To state our main results we need to introduce some notation. We begin by setting

$$\boldsymbol{\varphi} := (T, N, M, \theta, u, w),$$

$$\boldsymbol{\varphi}_h := (T_h, N_h, M_h, \theta_h, u_h, w_h),$$

$$\widehat{\boldsymbol{\varphi}}_h := (\widehat{T}_h, \widehat{N}_h, \widehat{M}_h, \widehat{\theta}_h, \widehat{u}_h, \widehat{w}_h),$$

$$\boldsymbol{u} := (u_1, u_2, u_3, u_4, u_5, u_6),$$

$$\boldsymbol{v} := (v_1, v_2, v_3, v_4, v_5, v_6),$$

where (T, N, M, θ, u, w) is the exact solution of (1.3) and (1.4), $(T_h,\ N_h,\ M_h,\ \theta_h,\ u_h,\ w_h)$ is the DG approximation defined by the weak formulation (2.2) and the numerical traces (4.2)-(2.5) where the functions C_{ij} are assumed to satisfy the conditions (3.1)-(3.4). The functions u_i and v_i are in V_h^k for some $k \geq 0$. We define the error of approximation as

$$e_\varphi = \varphi - \varphi_h, \qquad \widehat{e}_\varphi = \varphi - \widehat{\varphi}_h,$$

for any $\varphi \in \{T, N, M, \theta, u, w\}$, and set

$$\boldsymbol{e} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h, \qquad \widehat{\boldsymbol{e}} = \boldsymbol{\varphi} - \widehat{\boldsymbol{\varphi}}_h.$$

The error in the numerical traces of φ_h is defined as

$$\|\widehat{e}_\varphi\|_{L^\infty(\mathcal{E}_h)} := \max_{x_j \in \mathcal{E}_h} |\widehat{e}_\varphi(x_j)|,$$

and the global error in the numerical traces is set to be

$$\|\widehat{\boldsymbol{e}}\|_{L^\infty(\mathcal{E}_h)} := \max_{\varphi\in\{T,N,M,\theta,u,w\}} \|\widehat{e}_\varphi\|_{L^\infty(\mathcal{E}_h)}\,.$$

We define

$$|\boldsymbol{u}|^2_{\mathcal{A}_h} := \Theta_i(\boldsymbol{u}) + \Theta_j(\boldsymbol{u}), \tag{2.16}$$

where

$$\Theta_i(\boldsymbol{u}) = d^2(u_1,u_1)_{\Omega_h} + d^2(u_2,u_2)_{\Omega_h} + (u_3,u_3)_{\Omega_h},$$

and

$$\Theta_j(\boldsymbol{u}) = -\langle 1, C_{16}[\![u_1]\!]^2 + C_{25}[\![u_2]\!]^2 + C_{34}[\![u_3]\!]^2 + C_{43}[\![u_4]\!]^2 + C_{52}[\![u_5]\!]^2 + C_{61}[\![u_6]\!]^2\rangle_{\mathcal{E}_h}.$$

Since we can rewrite the discrete energy identity (4.15) of Proposition 2.1 as

$$|\boldsymbol{\varphi}_h|^2_{\mathcal{A}_h} = \Theta_{loads}(\boldsymbol{\varphi}_h) + \Theta_{bc}(\boldsymbol{\varphi}_h),$$

we call this seminorm, the *energy* seminorm. The estimate of the approximation error in this seminorm plays a fundamental role in our analysis.

Next, we define Green's functions for the problem under consideration. For any superindex $\star \in \{T, N, M, \theta, u, w\}$, and any point $y \in (0,1)$, we define $(G^\star_{T,y}, G^\star_{N,y}, G^\star_{M,y}, G^\star_{\theta,y}, G^\star_{u,y}, G^\star_{w,y})$ as the solution of

$$\begin{aligned}
-d\,G^\star_{w,y}/dx \quad -G^\star_{\theta,y} - \kappa G^\star_{u,y} \quad &= d^2 G^\star_{T,y},\\
-d\,G^\star_{u,y}/dx \qquad\quad + \kappa G^\star_{w,y} \quad &= d^2 G^\star_{N,y},\\
-d\,G^\star_{\theta,y}/dx \qquad\qquad\quad &= G^\star_{M,y},\\
-d\,G^\star_{M,y}/dx \qquad\qquad\quad &= G^\star_{T,y},\\
-d\,G^\star_{N,y}/dx \qquad\quad + \kappa G^\star_{T,y} \quad &= 0,\\
-d\,G^\star_{T,y}/dx \qquad\quad - \kappa G^\star_{N,y} \quad &= 0,
\end{aligned} \tag{2.17}$$

in $(0,y)\cup(y,1)$ that satisfies the boundary conditions

$$G^\star_{w,y} = G^\star_{u,y} = G^\star_{\theta,y} = 0 \quad \text{on } \partial\Omega, \tag{2.18}$$

and the jump conditions

$$\begin{aligned} [\![G^\star_{w,y}]\!](y) &= \delta_{\star T}, & [\![G^\star_{T,y}]\!](y) &= \delta_{\star w}, \\ [\![G^\star_{u,y}]\!](y) &= \delta_{\star N}, & [\![G^\star_{N,y}]\!](y) &= \delta_{\star u}, \\ [\![G^\star_{\theta,y}]\!](y) &= \delta_{\star M}, & [\![G^\star_{M,y}]\!](y) &= \delta_{\star \theta}. \end{aligned} \tag{2.19}$$

Here, $\delta_{ab} = 1$ if $a = b$ and $\delta_{ab} = 0$ otherwise. For simplicity, we denote

$$\boldsymbol{G}^\star_y := (G^\star_{T,y}, G^\star_{N,y}, G^\star_{M,y}, G^\star_{\theta,y}, G^\star_{u,y}, G^\star_{w,y}).$$

We also define, for $z \in \{0,1\}$,

$$\boldsymbol{G}^\star_z = \lim_{y\to z} \boldsymbol{G}^\star_y.$$

We denote by $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$ the usual norm and seminorm, respectively, in the Sobolev space $H^s(D)$ where D is any subset of Ω_h. We drop the subindex D whenever $D = \Omega_h$ or $D = \Omega$. We set

$$|\boldsymbol{u}|_{s,D} := (|u_1|^2_{s,D} + |u_2|^2_{s,D} + |u_3|^2_{s,D} + |u_4|^2_{s,D} + |u_5|^2_{s,D} + |u_6|^2_{s,D})^{1/2},$$

and

$$|\boldsymbol{G}|_{s,D} := \max_{x_j\in\mathcal{E}_h} \max_{\star\in\{T,N,M,\theta,u,w\}} |\boldsymbol{G}^\star_{x_j}|_{s,D}.$$

We are now ready to state and discuss our main results. In the rest of the paper, C denotes a generic constant which is not necessarily the same in each appearance, and it is independent of the meshsize h and the thickness parameter d even though we might not explicitly state it.

Theorem 2.3. *Let $k \geq 0$ be a polynomial degree and suppose that $\boldsymbol{\varphi}$ belongs to $[H^{k+1}(\Omega_h)]^6$. Let $\boldsymbol{\varphi}_h$ be the DG solution defined by the weak formulation* (2.2) *with $k_i = k$ for all $i = 1, \ldots, 6$, and the numerical traces* (4.2)-(2.5) *where the functions C_{ij} satisfy the conditions* (3.1)-(3.4)*. Then, for small enough h, we have that*

$$|\boldsymbol{e}|_{\mathcal{A}_h} \;\leq\; C\, h^{k+1/2} |\boldsymbol{\varphi}|_{k+1}, \tag{2.20}$$

and that

$$\|\boldsymbol{e}\|_0 \;\leq\; C\, h^{k+1} |\boldsymbol{\varphi}|_{k+1}, \tag{2.21}$$

for some constant C independent of h and d.

Theorem 2.4. *With the same hypotheses as those of Theorem 2.3 we have that*

$$\|\widehat{\boldsymbol{e}}\|_{L^\infty(\mathcal{E}_h)} \leq C\; h^{2k+1} |\boldsymbol{G}|_{k+1}\, |\boldsymbol{\varphi}|_{k+1}. \tag{2.22}$$

Note that all of the estimates appearing in the above theorems show that, the DG method under consideration is *locking-free* for any $k \geq 0$, because the constants appearing on the right-hand side of all the estimates are independent of the parameter d and because the seminorms appearing on the right-hand side of the estimates can be bounded uniformly with respect to d. See [20] for a detailed explanation of this in the context of Timoshenko beams. A similar remark is valid for the seminorms of the Green's functions, see [39].

Note also that the above results imply that the DG method converges with the optimal order of $k+1$ in the L^2-norm for *all* variables, and with order $k+1/2$ in the energy norm. They also imply that *all* the numerical traces superconverge with order $2\,k+1$ at each node. In Section 4.7, we verify that the error estimate in the energy seminorm is sharp. These results extend what has been done by Celiker *et al.* in [15, 16, 17, 20] for DG methods for Timoshenko beams.

2.3 Proofs

2.3.1 Sketch of proofs

In this subsection, we give a brief outline of the main steps of our proofs. We proceed in three steps. We begin with estimating the errors in the energy seminorm in terms of the errors in the L^2-norm.

Lemma 2.5. *We have*

$$|\boldsymbol{e}|_{\mathcal{A}_h} \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1} + Ch^{1/2}\,\|\boldsymbol{e}\|_0$$

for some constant C independent of h and d.

Next we show that the error in the numerical traces can be estimated in terms of the error in the L^2-norm and the seminorms of the Green's functions.

Lemma 2.6. *Let x_j be an arbitrary node in $\mathcal{E}_h$. If h is sufficiently small, then we have for any φ in $\{T, N, M, \theta, u, w\}$ that*

$$\|\widehat{\boldsymbol{e}}\|_{L^\infty(\mathcal{E}_h)} \leq Ch^k\,\|\boldsymbol{e}\|_0\,|\boldsymbol{G}|_{k+1}$$

for some constant C independent of h and d.

Finally, we obtain an auxiliary estimate of the error in the L^2-norm.

Lemma 2.7. *We have*

$$\|\boldsymbol{e}\|_0 \leq Ch^{k+1}|\boldsymbol{\varphi}|_{k+1} + Ch^{1/2}\,\|\boldsymbol{e}\|_0 + Ch^k|\boldsymbol{G}|_{k+1}\,\|\boldsymbol{e}\|_0$$

for some constant C independent of h and d.

The final estimates in the L^2-norm, namely (2.21), now follows if we assume that h is small enough. Using (2.21) in Lemma 2.5 yields (2.20). Similarly, inserting it into Lemma 2.6 we get (3.5).

2.3.2 The error equations

To prove the lemmas in the previous subsection, we rely, as expected, on the error equations, namely,

$$-(e_w, v_1')_{\Omega_h} + \langle \widehat{e}_w, [\![v_1]\!] \rangle_{\mathcal{E}_h} + (e_\theta, v_1)_{\Omega_h} + (\kappa e_u, v_1)_{\Omega_h} = d^2 (e_T, v_1)_{\Omega_h}, \tag{2.23a}$$

$$-(e_u, v_2')_{\Omega_h} + \langle \widehat{e}_u, [\![v_2]\!] \rangle_{\mathcal{E}_h} - (\kappa e_w, v_2)_{\Omega_h} = d^2 (e_N, v_2)_{\Omega_h}, \tag{2.23b}$$

$$-(e_\theta, v_3')_{\Omega_h} + \langle \widehat{e}_\theta, [\![v_3]\!] \rangle_{\mathcal{E}_h} = (e_M, v_3)_{\Omega_h}, \tag{2.23c}$$

$$-(e_M, v_4')_{\Omega_h} + \langle \widehat{e}_M, [\![v_4]\!] \rangle_{\mathcal{E}_h} = (e_T, v_4)_{\Omega_h}, \tag{2.23d}$$

$$-(e_N, v_5')_{\Omega_h} + \langle \widehat{e}_N, [\![v_5]\!] \rangle_{\mathcal{E}_h} - (\kappa e_T, v_5)_{\Omega_h} = 0, \tag{2.23e}$$

$$-(e_T, v_6')_{\Omega_h} + \langle \widehat{e}_T, [\![v_6]\!] \rangle_{\mathcal{E}_h} + (\kappa e_N, v_6)_{\Omega_h} = 0, \tag{2.23f}$$

for any $v_i \in V_h^k$, $i = 1, \ldots, 6$. They are easily obtained by noting that the exact solution $\boldsymbol{\varphi}$ also satisfies the DG formulation (2.2).

We use the following notation

$$\mathbf{P}\boldsymbol{e} := (\mathsf{P}e_T, \mathsf{P}e_N, \mathsf{P}e_M, \mathsf{P}e_\theta, \mathsf{P}e_u, \mathsf{P}e_w)$$

where P is the L^2-orthogonal projection into V_h^k. We also set

$$\boldsymbol{\xi} := \boldsymbol{e} - \mathbf{P}\boldsymbol{e} = (\xi_T, \xi_N, \xi_M, \xi_\theta, \xi_u, \xi_w) \tag{2.24}$$

where

$$\xi_\varphi := \varphi - \mathsf{P}\varphi$$

for $\varphi \in \{T, N, M, \theta, u, w\}$.

2.3.3 Proof of Lemma 2.5

We begin with expressing the DG method in classical mixed formulation. Inserting the definition of the numerical traces (4.2)-(2.5) into the weak formulation (2.2) and adding the resulting equations we get after some simple algebraic manipulations

$$\mathcal{A}_h(\boldsymbol{\varphi}_h; \boldsymbol{v}) = b_h(\boldsymbol{v})$$

where, writing $\mathcal{A}_h$ for $\mathcal{A}_h(\boldsymbol{u}; \boldsymbol{v})$, and similarly for $\mathcal{A}_1$, $\mathcal{A}_{2,i}$ etc.

$$\mathcal{A}_h(\boldsymbol{u}; \boldsymbol{v}) := \mathcal{A}_1(\boldsymbol{u}; \boldsymbol{v}) + \mathcal{A}_{2,i}(\boldsymbol{u}; \boldsymbol{v}) + \mathcal{A}_{2,\partial}(\boldsymbol{u}; \boldsymbol{v})$$

where

$$\begin{aligned}
\mathcal{A}_1 := &- (u_6, v_1')_{\Omega_h} + (u_3, v_1)_{\Omega_h} + (\kappa u_5, v_1)_{\Omega_h} - d^2 (u_1, v_1)_{\Omega_h} \\
&- (u_5, v_2')_{\Omega_h} - (\kappa u_6, v_2)_{\Omega_h} - d^2 (u_2, v_2)_{\Omega_h} \\
&- (u_4, v_3')_{\Omega_h} - (u_3, v_3)_{\Omega_h} \\
&- (u_3, v_4')_{\Omega_h} - (u_1, v_4)_{\Omega_h} \\
&- (u_2, v_5')_{\Omega_h} + (\kappa u_1, v_5)_{\Omega_h} \\
&- (u_1, v_6')_{\Omega_h} - (\kappa u_2, v_6)_{\Omega_h},
\end{aligned}$$

$$\mathcal{A}_{2,i} := \sum\nolimits_{j=1}^{6} \mathcal{A}_{2,i}^{(j)}, \qquad \mathcal{A}_{2,\partial} := \sum\nolimits_{j=1}^{6} \mathcal{A}_{2,\partial}^{(j)}$$

where, defining

$$\mathsf{C}_j := C_{j1}[\![u_6]\!] + C_{j2}[\![u_5]\!] + C_{j3}[\![u_4]\!] + C_{j4}[\![u_3]\!] + C_{j5}[\![u_2]\!] + C_{j6}[\![u_1]\!]$$

for $j = 1, \dots, 6$ we have

$$\mathcal{A}_{2,i}^{(1)} = \langle \{\!\{u_6\}\!\} + \mathsf{C}_1, [\![v_1]\!] \rangle_{\mathcal{E}_h^\circ}, \qquad \mathcal{A}_{2,i}^{(2)} = \langle \{\!\{u_5\}\!\} + \mathsf{C}_2, [\![v_2]\!] \rangle_{\mathcal{E}_h^\circ},$$

$$\mathcal{A}_{2,i}^{(3)} = \langle \{\!\{u_4\}\!\} + \mathsf{C}_3, [\![v_3]\!] \rangle_{\mathcal{E}_h^\circ}, \qquad \mathcal{A}_{2,i}^{(4)} = \langle \{\!\{u_3\}\!\} + \mathsf{C}_4, [\![v_4]\!] \rangle_{\mathcal{E}_h^\circ},$$

$$\mathcal{A}_{2,i}^{(5)} = \langle \{\!\{u_2\}\!\} + \mathsf{C}_5, [\![v_5]\!] \rangle_{\mathcal{E}_h^\circ}, \qquad \mathcal{A}_{2,i}^{(6)} = \langle \{\!\{u_1\}\!\} + \mathsf{C}_6, [\![v_6]\!] \rangle_{\mathcal{E}_h^\circ},$$

and

$$\mathcal{A}_{2,\partial}^{(1)} = -(u_3(0^+) - C_{41}(0)u_6(0^+) - C_{42}(0)u_5(0^+) - C_{43}(0)u_4(0^+))v_4(0^+),$$

$$\mathcal{A}_{2,\partial}^{(2)} = -(u_2(0^+) - C_{51}(0)u_6(0^+) - C_{52}(0)u_5(0^+) - C_{53}(0)u_4(0^+))v_5(0^+),$$

$$\mathcal{A}_{2,\partial}^{(3)} = -(u_1(0^+) - C_{61}(0)u_6(0^+) - C_{62}(0)u_5(0^+) - C_{63}(0)u_4(0^+))v_6(0^+),$$

$$\mathcal{A}_{2,\partial}^{(4)} = \quad (u_3(1^-) + C_{41}(1)u_6(1^-) + C_{42}(1)u_5(1^-) + C_{43}(1)u_4(1^-))v_4(1^-),$$

$$\mathcal{A}_{2,\partial}^{(5)} = \quad (u_2(1^-) + C_{51}(1)u_6(1^-) + C_{52}(1)u_5(1^-) + C_{53}(1)u_4(1^-))v_5(1^-),$$

$$\mathcal{A}_{2,\partial}^{(6)} = \quad (u_1(1^-) + C_{61}(1)u_6(1^-) + C_{62}(1)u_5(1^-) + C_{63}(1)u_4(1^-))v_6(1^-).$$

Finally, $b_h := b_1 + b_2$ where

$$b_1 = (p, v_5)_{\Omega_h} + (q, v_6)_{\Omega_h}$$

and

$$\begin{aligned} b_2 = \quad & w_0[v_1(0^+) + C_{61}(0)v_6(0^+) + C_{51}(0)v_5(0^+) + C_{41}(0)v_4(0^+)] \\ & + u_0[v_2(0^+) + C_{62}(0)v_6(0^+) + C_{52}(0)v_5(0^+) + C_{42}(0)v_4(0^+)] \\ & + \theta_0[v_3(0^+) + C_{63}(0)v_6(0^+) + C_{53}(0)v_5(0^+) + C_{43}(0)v_4(0^+)] \\ & - w_1[v_1(1^-) - C_{61}(1)v_6(1^-) - C_{51}(1)v_5(1^-) - C_{41}(1)v_4(1^-)] \\ & - u_1[v_2(1^-) - C_{62}(1)v_6(1^-) - C_{52}(1)v_5(1^-) - C_{42}(1)v_4(1^-)] \\ & - \theta_1[v_3(1^-) - C_{63}(1)v_6(1^-) - C_{53}(1)v_5(1^-) - C_{43}(1)v_4(1^-)]. \end{aligned}$$

With this notation we can write one of the main ingredients of our error analysis, namely the Galerkin orthogonality property, as

$$\mathcal{A}_h(\boldsymbol{e};\boldsymbol{v}) = 0 \qquad \text{for all} \quad \boldsymbol{v} \in [V_h^k]^6. \tag{2.25}$$

This follows immediately by adding the error equations (2.23). The second property we are going to use is

$$|\boldsymbol{v}|_{\mathcal{A}_h}^2 = -\mathcal{A}_h(\boldsymbol{v};\boldsymbol{v}) \qquad \text{for all} \quad \boldsymbol{v} \in [V_h^k]^6. \tag{2.26}$$

Lemma 2.5 follows from the following auxiliary results.

Lemma 2.8. *We have that* $|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}^2 = J_1 + J_{2,i} + J_{2,\partial}$ *where*

$$J_1 = (\kappa\xi_u, \mathsf{P}e_T)_{\Omega_h} - (\kappa\xi_w, \mathsf{P}e_N)_{\Omega_h} - (\kappa\xi_T, \mathsf{P}e_u)_{\Omega_h} + (\kappa\xi_N, \mathsf{P}e_w)_{\Omega_h},$$

and

$$J_{2,i} = \mathcal{A}_{2,i}(\boldsymbol{\xi};\mathbf{P}\boldsymbol{e}), \qquad J_{2,\partial} = \mathcal{A}_{2,\partial}(\boldsymbol{\xi};\mathbf{P}\boldsymbol{e}).$$

Lemma 2.9. *The following estimates hold*

$$|\boldsymbol{\xi}|_{\mathcal{A}_h} \leq C\, h^{k+1/2} |\boldsymbol{\varphi}|_{k+1}, \tag{2.27a}$$

$$|J_1| \leq C\, h^{k+1} |\boldsymbol{\varphi}|_{k+1} \left\|\boldsymbol{e}\right\|_0, \tag{2.27b}$$

$$|J_{2,i}| + |J_{2,\partial}| \leq C\, h^{k+1/2} |\boldsymbol{\varphi}|_{k+1} |\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}, \tag{2.27c}$$

for some constant C *independent of* h *and* d.

We prove these results in several steps.

Step 1: Proof of the auxiliary Lemma 2.8. We have

$$\begin{aligned}
|\mathbf{P}\boldsymbol{e}|^2_{\mathcal{A}_h} &= -\mathcal{A}_h(\mathbf{P}\boldsymbol{e};\mathbf{P}\boldsymbol{e}) \qquad \text{by (2.26)},\\
&= -\mathcal{A}_h(\boldsymbol{e}-\boldsymbol{\xi};\mathbf{P}\boldsymbol{e}) \quad \text{by (2.24)},\\
&= \mathcal{A}_h(\boldsymbol{\xi};\mathbf{P}\boldsymbol{e}) \qquad \text{by (2.25)},\\
&= J_1 + J_{2,i} + J_{2,\partial},
\end{aligned}$$

by the orthogonality properties of the L^2-projection operator P. This finishes the proof of Lemma 2.8.

Note that for arches with piecewise constant κ on oh, the term J_1 vanishes by the orthogonality properties of P.

Step 2: Estimate of $|\boldsymbol{\xi}|_{\mathcal{A}_h}$. We will need the following lemma which contains the approximation properties of P which can be found, for example, in [22].

Lemma 2.10. *Let $I_j \subset \Omega_h$ be an arbitrary element, and suppose that $\phi \in H^{t+1}(I_j)$ for some non-negative real number t. Then*

$$\begin{aligned}
\|\phi - \mathsf{P}\phi\|_{0,I_j} &\leq C|\phi|_{\sigma+1,I_j}h_j^{\sigma+1},\\
|(\phi-\mathsf{P}\phi)(x_{j-1}^+)| + |(\phi-\mathsf{P}\phi)(x_j^-)| &\leq C|\phi|_{\sigma+1,I_j}h_j^{\sigma+1/2},
\end{aligned}$$

for any $0 \leq \sigma \leq \min(k,t)$, and for some constant C depending solely on t.

We have that

$$|\boldsymbol{\xi}|^2_{\mathcal{A}_h} = \Theta_i(\boldsymbol{\xi}) + \Theta_j(\boldsymbol{\xi}) \tag{2.28}$$

where

$$\Theta_i(\boldsymbol{\xi}) = d^2(\xi_T,\xi_T)_{\Omega_h} + d^2(\xi_N,\xi_N)_{\Omega_h} + (\xi_M,\xi_M)_{\Omega_h},$$

and

$$\begin{aligned}\Theta_j(\boldsymbol{\xi}) &= -\langle 1, C_{16}[\![\xi_T]\!]^2 + C_{25}[\![\xi_N]\!]^2 + C_{34}[\![\xi_M]\!]^2 + C_{43}[\![\xi_\theta]\!]^2 + C_{52}[\![\xi_u]\!]^2 + C_{61}[\![\xi_w]\!]^2\rangle_{\mathcal{E}_h} \\ &= \mathsf{c}\langle 1, [\![\xi_T]\!]^2 + [\![\xi_N]\!]^2 + [\![\xi_M]\!]^2 + [\![\xi_\theta]\!]^2 + [\![\xi_u]\!]^2 + [\![\xi_w]\!]^2\rangle_{\mathcal{E}_h},\end{aligned}$$

by assumption (3.1).

Now, since $d < 1$ we have

$$\begin{aligned}\Theta_i(\boldsymbol{\xi}) &\le \|\xi_T\|_0^2 + \|\xi_N\|_0^2 + \|\xi_M\|_0^2 \\ &\le Ch^{2k+2}(|T|_{k+1}^2 + |N|_{k+1}^2 + |M|_{k+1}^2) \\ &\le Ch^{2k+2}|\boldsymbol{\varphi}|_{k+1}^2\end{aligned} \tag{2.29}$$

by the approximation properties of the previous lemma with $\sigma = k$.

Next we estimate $\Theta_j(\boldsymbol{\xi})$. By the approximation properties of P

$$\mathsf{c}\langle 1, [\![\xi_\varphi]\!]^2\rangle_{\mathcal{E}_h} \le C\, h^{2k+1}|\boldsymbol{\varphi}|_{k+1}^2$$

for all $\varphi \in \{T, N, M, \theta, u, w\}$, where we have absorbed c in C since it is a constant of order one. Hence, we get

$$\Theta_j(\boldsymbol{\xi}) \le C\, h^{2k+1}|\boldsymbol{\varphi}|_{k+1}^2. \tag{2.30}$$

Inserting the estimates (2.29) and (2.30) into (2.28), and taking the square root of both sides of the resulting estimate yields (2.27a).

Step 3: Estimate of J_1. We only show how to estimate one of the terms appearing in J_1,

the remaining three terms can be estimated in a similar fashion. We proceed as follows

$$
\begin{aligned}
|(\kappa\xi_u, \mathsf{P}e_T)_{\Omega_h}| &\leq \max_{x\in\Omega} |\kappa(x)| \, \|\xi_u\|_0 \, \|\mathsf{P}e_T\|_0 && \text{by Cauchy-Schwarz inequality,} \\
&\leq Ch^{k+1}|\boldsymbol{\varphi}|_{k+1} \, \|\mathsf{P}e_T\|_0 && \text{by Lemma 2.10,} \\
&\leq Ch^{k+1}|\boldsymbol{\varphi}|_{k+1} \, \|e_T\|_0 && \text{by the continuity of } \mathsf{P}, \\
&\leq Ch^{k+1}|\boldsymbol{\varphi}|_{k+1} \, \|\boldsymbol{e}\|_0 \, .
\end{aligned}
$$

This finishes the proof of (2.27b).

Step 4: Estimate of $J_{2,i}$. To estimate $J_{2,i}$ we note that

$$
\begin{aligned}
\langle \{\!\{\xi_w\}\!\}, [\![\mathsf{P}e_T]\!] \rangle_{\mathcal{E}_h^\circ} &\leq \langle 1/\mathsf{c}, \{\!\{\xi_w\}\!\}^2 \rangle_{\mathcal{E}_h^\circ}^{1/2} \, \langle \mathsf{c}, [\![\mathsf{P}e_T]\!]^2 \rangle_{\mathcal{E}_h^\circ}^{1/2} \\
&\leq Ch^{k+1/2}|w|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h} \\
&\leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}
\end{aligned}
$$

where we have used the Cauchy-Schwarz inequality and the approximation properties of P. We also have that

$$
\begin{aligned}
\langle C_{11}[\![\xi_w]\!], [\![\mathsf{P}e_T]\!] \rangle_{\mathcal{E}_h^\circ} &\leq \langle C_{11}^2/\mathsf{c}, [\![\xi_w]\!]^2 \rangle_{\mathcal{E}_h^\circ}^{1/2} \, \langle \mathsf{c}, [\![\mathsf{P}e_T]\!]^2 \rangle_{\mathcal{E}_h^\circ}^{1/2} \\
&\leq Ch^{k+1/2}|w|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h} \\
&\leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}
\end{aligned}
$$

where we have used the assumption (3.3). Similarly, we obtain

$$\langle C_{12}[\![\xi_u]\!], [\![\mathsf{P}e_T]\!]\rangle_{\mathcal{E}_h^\circ} \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

$$\langle C_{13}[\![\xi_\theta]\!], [\![\mathsf{P}e_T]\!]\rangle_{\mathcal{E}_h^\circ} \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

$$\langle C_{14}[\![\xi_M]\!], [\![\mathsf{P}e_T]\!]\rangle_{\mathcal{E}_h^\circ} \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

$$\langle C_{15}[\![\xi_N]\!], [\![\mathsf{P}e_T]\!]\rangle_{\mathcal{E}_h^\circ} \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

$$\langle C_{16}[\![\xi_T]\!], [\![\mathsf{P}e_T]\!]\rangle_{\mathcal{E}_h^\circ} \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}.$$

Collecting these estimates we get that

$$|J_{2,i}^{(1)}| = |\mathcal{A}_{2,i}^{(1)}(\boldsymbol{\xi}; \mathbf{P}\boldsymbol{e})| \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}.$$

Following the same steps, we can prove that

$$|J_{2,i}^{(\ell)}| = |\mathcal{A}_{2,i}^{(\ell)}(\boldsymbol{\xi}; \mathbf{P}\boldsymbol{e})| \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

for all $\ell = 2, \ldots, 6$. Since $J_{2,i} = \sum_{\ell=1}^{6} J_{2,i}^{(\ell)}$ we get

$$|J_{2,i}| \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}. \tag{2.31}$$

Step 5: Estimate of $J_{2,\partial}$. The estimate $J_{2,\partial}$ follows similar lines as those of the estimate of $J_{2,i}$. Thus, by the approximation properties of P we have

$$\begin{aligned}
|\xi_M(0^+)\mathsf{P}e_\theta(0^+)| &= \frac{1}{\mathsf{c}}|\xi_M(0^+)||\mathsf{c}\,\mathsf{P}e_\theta(0^+)| \\
&\leq Ch^{k+1/2}|M|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}, \\
&\leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},
\end{aligned}$$

and

$$\begin{aligned}
|C_{41}(0)\xi_w(0^+)\mathsf{P}e_\theta(0^+)| &= \frac{|C_{41}(0)|}{\mathsf{c}}|\xi_w(0^+)||\mathsf{c}\,\mathsf{P}e_\theta(0^+)| \\
&\leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}.
\end{aligned}$$

Similarly,

$$|C_{42}(0)\xi_u(0^+)\mathsf{P}e_\theta(0^+)| \le Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

$$|C_{43}(0)\xi_\theta(0^+)\mathsf{P}e_\theta(0^+)| \le Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

and hence

$$|J_{2,\partial}^{(1)}| = |\mathcal{A}_{2,\partial}^{(1)}(\boldsymbol{\xi};\mathbf{P}\boldsymbol{e})| \le Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}.$$

Following the same steps, we can prove that

$$|J_{2,\partial}^{(\ell)}| = |\mathcal{A}_{2,\partial}^{(\ell)}(\boldsymbol{\xi};\mathbf{P}\boldsymbol{e})| \le Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h},$$

for all $\ell = 2,\dots,6$. Since $J_{2,\partial} = \sum_{\ell=1}^{6} J_{2,\partial}^{(\ell)}$ we get

$$|J_{2,\partial}| \le Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}. \tag{2.32}$$

The estimate (2.27c) follows from (2.31) and (2.32).

Step 6: Proof of Lemma 2.5. By inserting the estimates (2.27b) and (2.27c) into the expression for $|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}^2$ in Lemma 2.8 we get

$$|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}^2 \le Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h} + Ch^{k+1}|\boldsymbol{\varphi}|_{k+1}\left\|\boldsymbol{e}\right\|_0.$$

Applying the Young's inequality to the first term on the right-hand side we get

$$Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h} \le Ch^{2k+1}|\boldsymbol{\varphi}|_{k+1}^2 + \frac{1}{2}|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}^2,$$

and hence

$$|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h}^2 \le Ch^{2k+1}|\boldsymbol{\varphi}|_{k+1}^2 + Ch^{k+1}|\boldsymbol{\varphi}|_{k+1}\left\|\boldsymbol{e}\right\|_0.$$

Applying the Young's inequality once more gives

$$Ch^{k+1}|\boldsymbol{\varphi}|_{k+1}\left\|\boldsymbol{e}\right\|_0 \le Ch^{2k+1}|\boldsymbol{\varphi}|_{k+1}^2 + \frac{h}{2}\left\|\boldsymbol{e}\right\|_0^2.$$

Thus,

$$|\mathbf{P}\boldsymbol{e}|^2_{\mathcal{A}_h} \leq Ch^{2k+1}|\boldsymbol{\varphi}|^2_{k+1} + Ch\,\|\boldsymbol{e}\|^2_0\,,$$

and hence

$$|\mathbf{P}\boldsymbol{e}|_{\mathcal{A}_h} \leq Ch^{k+1/2}|\boldsymbol{\varphi}|_{k+1} + Ch^{1/2}\,\|\boldsymbol{e}\|_0\,.$$

Combining this estimate with (2.27a) and applying the triangle inequality finishes the proof of Lemma 2.5

2.3.4 Proof of Lemma 2.6

To prove Lemma 2.6, we proceed in two steps.

Step 1: The error representation formulas. Our next result contains a representation formula for the errors in the numerical traces in terms of certain integrals involving the Green's functions. To state it, we need to introduce a projection operator. For any $\phi \in H^1(\Omega_h)$, the function $\Pi^+\phi \in V_h^k$ is defined on the element I_j by

$$(\phi - \Pi^+\phi, v)_{I_j} = 0 \qquad \forall v \in P^{k-1}(I_j),\ \text{if } k > 0, \tag{2.33a}$$

$$(\Pi^+\phi)(x^+_{j-1}) = \phi(x^+_{j-1}). \tag{2.33b}$$

Lemma 2.11 (Error representation formulas)**.** *Let $x_j \in \mathcal{E}_h$ be an arbitrary node and let G^{φ}_{w,x_j},G^{φ}_{u,x_j}, G^{φ}_{θ,x_j}, G^{φ}_{M,x_j}, G^{φ}_{N,x_j}, G^{φ}_{T,x_j}, for $\varphi \in \{T, N, M, \theta, u, w\}$, be the Green's functions defined by equations* (4.12)*,* (4.13) *and* (4.14)*. Then*

$$\widehat{e}_{\varphi}(x_j) = \Gamma^{\varphi}_{j,1} + \Gamma^{\varphi}_{j,2} + \Gamma^{\varphi}_{j,3},$$

where

$$\begin{aligned}\Gamma^{\varphi}_{j,1} := \quad & (e_w, (\Pi G^{\varphi}_{T,x_j} - G^{\varphi}_{T,x_j})')_{\Omega_h} \ + (e_u, (\Pi G^{\varphi}_{N,x_j} - G^{\varphi}_{N,x_j})')_{\Omega_h} \\ & +(e_\theta, (\Pi G^{\varphi}_{M,x_j} - G^{\varphi}_{M,x_j})')_{\Omega_h} + (e_M, (\Pi G^{\varphi}_{\theta,x_j} - G^{\varphi}_{\theta,x_j})')_{\Omega_h} \\ & +(e_N, (\Pi G^{\varphi}_{u,x_j} - G^{\varphi}_{u,x_j})')_{\Omega_h} \ + (e_T, (\Pi G^{\varphi}_{w,x_j} - G^{\varphi}_{w,x_j})')_{\Omega_h},\end{aligned}$$

$$\begin{aligned}\Gamma^{\varphi}_{j,2} := \quad & \sum_{i=1}^{\mathcal{N}} \widehat{e}_w(x_i)[G^{\varphi}_{T,x_j} - \Pi^+ G^{\varphi}_{T,x_j}](x_i^-) \ + \sum_{i=1}^{\mathcal{N}} \widehat{e}_u(x_i)[G^{\varphi}_{N,x_j} - \Pi^+ G^{\varphi}_{N,x_j}](x_i^-) \\ & + \sum_{i=1}^{\mathcal{N}} \widehat{e}_\theta(x_i)[G^{\varphi}_{M,x_j} - \Pi^+ G^{\varphi}_{M,x_j}](x_i^-) + \sum_{i=1}^{\mathcal{N}} \widehat{e}_M(x_i)[G^{\varphi}_{\theta,x_j} - \Pi^+ G^{\varphi}_{\theta,x_j}](x_i^-) \\ & + \sum_{i=1}^{\mathcal{N}} \widehat{e}_N(x_i)[G^{\varphi}_{u,x_j} - \Pi^+ G^{\varphi}_{u,x_j}](x_i^-) \ + \sum_{i=1}^{\mathcal{N}} \widehat{e}_T(x_i)[G^{\varphi}_{w,x_j} - \Pi^+ G^{\varphi}_{w,x_j}](x_i^-),\end{aligned}$$

and

$$\begin{aligned}\Gamma^{\varphi}_{j,3} := \quad & -(e_\theta + \kappa e_u - d^2 e_T, \Pi G^{\varphi}_{T,x_j} - G^{\varphi}_{T,x_j})_{\Omega_h} + (e_M, \Pi G^{\varphi}_{M,x_j} - G^{\varphi}_{M,x_j})_{\Omega_h} \\ & +(\kappa e_w + d^2 e_N, \Pi G^{\varphi}_{N,x_j} - G^{\varphi}_{N,x_j})_{\Omega_h} \qquad + (e_T, \Pi G^{\varphi}_{\theta,x_j} - G^{\varphi}_{\theta,x_j})_{\Omega_h} \\ & +(\kappa e_T, \Pi G^{\varphi}_{u,x_j} - G^{\varphi}_{u,x_j})_{\Omega_h} \qquad - (\kappa e_N, \Pi G^{\varphi}_{w,x_j} - G^{\varphi}_{w,x_j})_{\Omega_h}.\end{aligned}$$

To prove this lemma we need an auxiliary result which establishes a relation between the errors in the numerical traces and the Green's functions.

Lemma 2.12. *With the same notation as in Lemma 2.11 set*

$$\begin{aligned}\Theta^{\varphi}_j := \quad & \langle \widehat{e}_w, [\![G^{\varphi}_{T,x_j}]\!] \rangle_{\varepsilon_h} + \langle \widehat{e}_u, [\![G^{\varphi}_{N,x_j}]\!] \rangle_{\varepsilon_h} \\ & +\langle \widehat{e}_\theta, [\![G^{\varphi}_{M,x_j}]\!] \rangle_{\varepsilon_h} + \langle \widehat{e}_M, [\![G^{\varphi}_{\theta,x_j}]\!] \rangle_{\varepsilon_h} \\ & +\langle \widehat{e}_N, [\![G^{\varphi}_{u,x_j}]\!] \rangle_{\varepsilon_h} + \langle \widehat{e}_T, [\![G^{\varphi}_{w,x_j}]\!] \rangle_{\varepsilon_h}.\end{aligned}$$

Then, we have

$$\Theta^{\varphi}_j = \Lambda^{\varphi}_{j,1} + \Lambda^{\varphi}_{j,2} + \Lambda^{\varphi}_{j,2},$$

where

$$\begin{aligned}\Lambda^{\varphi}_{j,1} := \quad & (e_w, (v_1 - G^{\varphi}_{T,x_j})')_{\Omega_h} + (e_u, (v_2 - G^{\varphi}_{N,x_j})')_{\Omega_h} \\ & + (e_\theta, (v_3 - G^{\varphi}_{M,x_j})')_{\Omega_h} + (e_M, (v_4 - G^{\varphi}_{\theta,x_j})')_{\Omega_h} \\ & + (e_N, (v_5 - G^{\varphi}_{u,x_j})')_{\Omega_h} + (e_T, (v_6 - G^{\varphi}_{w,x_j})')_{\Omega_h},\end{aligned}$$

$$\begin{aligned}\Lambda^{\varphi}_{j,2} := \quad & \langle \widehat{e}_w, [\![G^{\varphi}_{T,x_j} - v_1]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_u, [\![G^{\varphi}_{N,x_j} - v_2]\!] \rangle_{\mathcal{E}_h} \\ & + \langle \widehat{e}_\theta, [\![G^{\varphi}_{M,x_j} - v_3]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_M, [\![G^{\varphi}_{\theta,x_j} - v_4]\!] \rangle_{\mathcal{E}_h} \\ & + \langle \widehat{e}_N, [\![G^{\varphi}_{u,x_j} - v_5]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_T, [\![G^{\varphi}_{w,x_j} - v_6]\!] \rangle_{\mathcal{E}_h},\end{aligned}$$

and

$$\begin{aligned}\Lambda^{\varphi}_{j,3} := \quad & -(e_\theta + \kappa e_u - d^2 e_T, v_1 - G^{\varphi}_{T,x_j})_{\Omega_h} + (e_M, v_3 - G^{\varphi}_{M,x_j})_{\Omega_h} \\ & + (\kappa e_w + d^2 e_N, v_2 - G^{\varphi}_{N,x_j})_{\Omega_h} \qquad + (e_T, v_4 - G^{\varphi}_{\theta,x_j})_{\Omega_h} \\ & + (\kappa e_T, v_5 - G^{\varphi}_{u,x_j})_{\Omega_h} \qquad - (\kappa e_N, v_6 - G^{\varphi}_{w,x_j})_{\Omega_h}\end{aligned}$$

for all v_i in V_h^k for $i = 1, \ldots, 6$.

Proof. Since we can write $\Theta^{\varphi}_j = \Upsilon^{\varphi}_j + \Lambda^{\varphi}_{j,2}$ where

$$\begin{aligned}\Upsilon^{\varphi}_j := \quad & \langle \widehat{e}_w, [\![v_1]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_u, [\![v_2]\!] \rangle_{\mathcal{E}_h} \\ & + \langle \widehat{e}_\theta, [\![v_3]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_M, [\![v_4]\!] \rangle_{\mathcal{E}_h} \\ & + \langle \widehat{e}_N, [\![v_5]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_T, [\![v_6]\!] \rangle_{\mathcal{E}_h},\end{aligned}$$

and

$$\begin{aligned}\Delta_j^{\varphi} := \quad & \langle \widehat{e}_w, [\![G^{\varphi}_{T,x_j} - v_1]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_u, [\![G^{\varphi}_{N,x_j} - v_2]\!] \rangle_{\mathcal{E}_h} \\ & + \langle \widehat{e}_\theta, [\![G^{\varphi}_{M,x_j} - v_3]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_M, [\![G^{\varphi}_{\theta,x_j} - v_4]\!] \rangle_{\mathcal{E}_h} \\ & + \langle \widehat{e}_N, [\![G^{\varphi}_{u,x_j} - v_5]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_T, [\![G^{\varphi}_{w,x_j} - v_6]\!] \rangle_{\mathcal{E}_h}\end{aligned}$$

we only have to prove that

$$\Upsilon_j^{\varphi} = \Lambda_{j,1}^{\varphi} + \Lambda_{j,3}^{\varphi}. \tag{2.34}$$

To achieve this, we proceed as follows. First, note that, by the definition of the Green's functions (4.12), we have

$$\begin{aligned}
-(e_w, (G^{\varphi}_{T,x_j})')_{\Omega_h} - (\kappa e_w, G^{\varphi}_{N,x_j})_{\Omega_h} &= 0, \\
-(e_u, (G^{\varphi}_{N,x_j})')_{\Omega_h} + (\kappa e_u, G^{\varphi}_{T,x_j})_{\Omega_h} &= 0, \\
-(e_\theta, (G^{\varphi}_{M,x_j})')_{\Omega_h} &= -(e_\theta, G^{\varphi}_{T,x_j})_{\Omega_h}, \\
-(e_M, (G^{\varphi}_{\theta,x_j})')_{\Omega_h} &= (e_M, G^{\varphi}_{M,x_j})_{\Omega_h}, \\
-(e_N, (G^{\varphi}_{u,x_j})')_{\Omega_h} + (\kappa e_N, G^{\varphi}_{w,x_j})_{\Omega_h} &= d^2(e_N, G^{\varphi}_{N,x_j})_{\Omega_h}, \\
-(e_T, (G^{\varphi}_{w,x_j})')_{\Omega_h} - (\kappa e_T, G^{\varphi}_{\theta,x_j})_{\Omega_h} &= d^2(e_T, G^{\varphi}_{N,x_j})_{\Omega_h} + (e_T, G^{\varphi}_{\theta,x_j})_{\Omega_h}.
\end{aligned} \tag{2.35}$$

Adding all the error equations (2.23) we obtain

$$\begin{aligned}
\Upsilon_i^{\varphi} = \quad & (e_w, v_1')_{\Omega_h} - (e_\theta, v_1)_{\Omega_h} - (\kappa e_u, v_1)_{\Omega_h} + d^2(e_T, v_1)_{\Omega_h} \\
& + (e_u, v_2')_{\Omega_h} + (\kappa e_w, v_2)_{\Omega_h} + d^2(e_N, v_2)_{\Omega_h} \\
& + (e_\theta, v_3')_{\Omega_h} + (e_M, v_3)_{\Omega_h} \\
& + (e_M, v_4')_{\Omega_h} + (e_T, v_4)_{\Omega_h} \\
& + (e_N, v_5')_{\Omega_h} + (\kappa e_T, v_5)_{\Omega_h} \\
& + (e_T, v_6')_{\Omega_h} - (\kappa e_N, v_6)_{\Omega_h}.
\end{aligned} \tag{2.36}$$

Collecting all the terms in (2.35) on the left-hand side, adding the resulting equations, and then subtracting the result from (2.36), we reach at (2.34) by a simple regrouping of like terms. This completes the proof. □

We are now ready to prove Lemma 2.11.

Proof of Lemma 2.11. We begin by noting that, by the definition of the Green's functions, (4.13) and (4.14), we have

$$\Theta_j^\varphi = \widehat{e}_\varphi(x_j).$$

On the other hand, setting

$$(v_1, v_2, v_3, v_4, v_5, v_6) = (\Pi^+ G^\varphi_{T,x_j}, \Pi^+ G^\varphi_{N,x_j}, \Pi^+ G^\varphi_{M,x_j}, \Pi^+ G^\varphi_{\theta,x_j}, \Pi^+ G^\varphi_{u,x_j}, \Pi^+ G^\varphi_{w,x_j})$$

in Lemma 2.12, we get

$$\widehat{e}_\varphi(x_j) = \Lambda^\varphi_{j,1} + \Phi^\varphi_{j,2} + \Lambda^\varphi_{j,3}$$

where

$$\begin{aligned}
\Phi^\varphi_{j,2} := \quad & \langle \widehat{e}_w, [\![G^\varphi_{T,x_j} - \Pi^+ G^\varphi_{T,x_j}]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_u, [\![G^\varphi_{N,x_j} - \Pi^+ G^\varphi_{N,x_j}]\!] \rangle_{\mathcal{E}_h} \\
& + \langle \widehat{e}_\theta, [\![G^\varphi_{M,x_j} - \Pi^+ G^\varphi_{M,x_j}]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_M, [\![G^\varphi_{\theta,x_j} - \Pi^+ G^\varphi_{\theta,x_j}]\!] \rangle_{\mathcal{E}_h} \\
& + \langle \widehat{e}_N, [\![G^\varphi_{u,x_j} - \Pi^+ G^\varphi_{u,x_j}]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_T, [\![G^\varphi_{w,x_j} - \Pi^+ G^\varphi_{w,x_j}]\!] \rangle_{\mathcal{E}_h}.
\end{aligned}$$

Note, by (2.33b), that

$$(G^\varphi_{\star,x_j} - \Pi^+ G^\varphi_{\star,x_j})(x_i^+) = 0 \quad \text{for } i = 0, 1, \ldots, \mathcal{N} - 1$$

for any $\star \in \{T, N, M, \theta, u, w\}$. Hence, we have $\Phi^\varphi_{j,2} = \Lambda^\varphi_{j,2}$. This completes the proof. □

Step 2: Estimating the error in the numerical traces. Here, we apply the approximation properties of the projection operator Π^+ to the error representation formulas of Lemma 2.11 to prove Lemma 2.6. For a proof of the following lemma see [34] and, for example, [14] or [28].

Lemma 2.13. *Let $I_j \subset \Omega_h$ be an arbitrary element, and suppose that $\phi \in H^{t+1}(I_j)$ for some non-negative real number t. Then*

$$\begin{aligned} \left\|\phi - \Pi^+\phi\right\|_{0,I_j} &\leq C|\phi|_{\sigma+1,I_j} h_j^{\sigma+1}, \\ \left\|(\phi - \Pi^+\phi)'\right\|_{0,I_j} &\leq C|\phi|_{\sigma+1,I_j} h_j^{\sigma}, \\ |(\phi - \Pi^+\phi)(x_j^-)| &\leq C|\phi|_{\sigma+1,I_j} h_j^{\sigma+1/2}, \end{aligned}$$

for any $0 \leq \sigma \leq \min(k,t)$, and for some constant C depending solely on t.

Proof of Lemma 2.6. The estimate follows by estimating each one of the terms appearing on the right-hand side of the expression for $\widehat{e}_\varphi(x_j)$ given in Lemma 2.11. We only show how to estimate three typical terms (one term from each of $\Lambda_{j,1}^\varphi$, $\Lambda_{j,2}^\varphi$, and $\Lambda_{j,3}^\varphi$) since the estimation of the remaining terms are similar. By Cauchy-Schwarz inequality and the approximation properties of Π^+ given in Lemma 2.13 we have

$$\begin{aligned} |(e_w, (\Pi^+ G_{T,x_j}^\varphi - G_{T,x_j}^\varphi)')_{\Omega_h}| &\leq \|e_w\|_0 \left\|(\Pi^+ G_{T,x_j}^\varphi - G_{T,x_j}^\varphi)'\right\|_0 \\ &\leq \|\boldsymbol{e}\|_0 \cdot C\, h^k\, |G_{T,x_j}^\varphi|_{k+1} \\ &\leq Ch^k \|\boldsymbol{e}\|_0 |\boldsymbol{G}|_{k+1}. \end{aligned}$$

Moving the maximum of $|\widehat{e}_w(x_i)|$ over $i = 1, \ldots, \mathcal{N}$ outside the summation we get

$$\begin{aligned}\Big| \sum_{i=1}^{\mathcal{N}} \widehat{e}_w(x_i)[G^{\varphi}_{T,x_j} - \Pi^+ G^{\varphi}_{T,x_j}](x_i^-) \Big| &\leq \|\widehat{e}_w\|_{L^\infty(\mathcal{E}_h)} \sum_{i=1}^{\mathcal{N}} |(G^{\varphi}_{T,x_j} - \Pi^+ G^{\varphi}_{T,x_j})(x_i^-)| \\ &\leq \|\widehat{e}_w\|_{L^\infty(\mathcal{E}_h)} \cdot Ch^{k+1/2} |G^{\varphi}_{T,x_j}|_{k+1} \\ &\leq Ch^{k+1/2} \|\widehat{\boldsymbol{e}}\|_{L^\infty(\mathcal{E}_h)} |\boldsymbol{G}|_{k+1}.\end{aligned}$$

Since $d < 1$, and κ is bounded

$$\begin{aligned}|(\kappa e_w + d^2 e_N, \Pi^+ G^{\varphi}_{N,x_j} - G^{\varphi}_{N,x_j})_{\Omega_h}| &\leq \big\|\kappa e_w + d^2 e_N\big\|_0 \big\|\Pi^+ G^{\varphi}_{N,x_j} - G^{\varphi}_{N,x_j}\big\|_0 \\ &\leq (C \|e_w\|_0 + \|e_N\|_0) \big\|\Pi^+ G^{\varphi}_{N,x_j} - G^{\varphi}_{N,x_j}\big\|_0 \\ &\leq C \|\boldsymbol{e}\|_0 \cdot C\, h^{k+1} |G^{\varphi}_{N,x_j}|_{k+1} \\ &\leq Ch^{k+1} \|\boldsymbol{e}\|_0 |\boldsymbol{G}|_{k+1}.\end{aligned}$$

Estimating the remaining terms similarly, and collecting the resulting estimates we obtain

$$|\widehat{e}_\varphi(x_j)| \leq Ch^k \|\boldsymbol{e}\|_0 |\boldsymbol{G}|_{k+1} + Ch^{k+1/2} \|\widehat{\boldsymbol{e}}\|_{L^\infty(\mathcal{E}_h)} |\boldsymbol{G}|_{k+1}.$$

Note that the right-hand side of this estimate does not depend on x_j or φ. Hence, taking the maximum of both sides over $x_j \in \mathcal{E}_h$ and $\varphi \in \{T, N, M, \theta, u, w\}$ we get

$$\|\widehat{\boldsymbol{e}}\|_{L^\infty(\mathcal{E}_h)} \leq Ch^k \|\boldsymbol{e}\|_0 |\boldsymbol{G}|_{k+1} + Ch^{k+1/2} \|\widehat{\boldsymbol{e}}\|_{L^\infty(\mathcal{E}_h)} |\boldsymbol{G}|_{k+1}.$$

Assuming that h is small enough so that

$$Ch^{k+1/2} |\boldsymbol{G}|_{k+1} \leq \alpha < 1$$

we reach at the desired estimate. □

2.3.5 Proof of Lemma 2.7

To prove Lemma 2.7 we proceed in several steps.

Step 1: The representation formulas. The following lemma is an auxiliary result which contains suitable expressions for (e_φ, ψ) for $\varphi \in \{T, N, M, \theta, u, w\}$ where ψ is an arbitrary function in $L^2(\Omega_h)$. We use the notation $\xi_\phi^+ := \phi - \Pi^+\phi$ where Π^+ is the projection operator defined by (2.33).

Lemma 2.14. *Let $\psi \in L^2(\Omega_h)$ and let $\widetilde{\Psi}(x) := \int_0^x \psi(s)ds$. Define Ψ as the function on Ω_h whose restriction to the element $I_j = (x_{j-1}, x_j) \in \Omega_h$ is*

$$\Psi|_{I_j}(x) = \widetilde{\Psi}(x) - \widetilde{\Psi}(x_{j-1}).$$

Then the following expressions hold

$$(e_w, \psi)_{\Omega_h} = -((\xi_w^+)', \xi_\Psi^+)_{\Omega_h} + (\mathcal{R} - \mathcal{S})(w) + (e_\theta + \kappa e_u - d^2 e_T, \Pi^+\Psi)_{\Omega_h}, \tag{2.37a}$$

$$(e_u, \psi)_{\Omega_h} = -((\xi_u^+)', \xi_\Psi^+)_{\Omega_h} + (\mathcal{R} - \mathcal{S})(u) - (\kappa e_w + d^2 e_N, \Pi^+\Psi)_{\Omega_h}, \tag{2.37b}$$

$$(e_\theta, \psi)_{\Omega_h} = -((\xi_\theta^+)', \xi_\Psi^+)_{\Omega_h} + (\mathcal{R} - \mathcal{S})(\theta) - (e_M, \Pi^+\Psi)_{\Omega_h}, \tag{2.37c}$$

$$(e_M, \psi)_{\Omega_h} = -((\xi_M^+)', \xi_\Psi^+)_{\Omega_h} + (\mathcal{R} - \mathcal{S})(M) - (e_T, \Pi^+\Psi)_{\Omega_h}, \tag{2.37d}$$

$$(e_N, \psi)_{\Omega_h} = -((\xi_N^+)', \xi_\Psi^+)_{\Omega_h} + (\mathcal{R} - \mathcal{S})(N) - (\kappa e_T, \Pi^+\Psi)_{\Omega_h}, \tag{2.37e}$$

$$(e_T, \psi)_{\Omega_h} = -((\xi_T^+)', \xi_\Psi^+)_{\Omega_h} + (\mathcal{R} - \mathcal{S})(T) + (\kappa e_N, \Pi^+\Psi)_{\Omega_h}, \tag{2.37f}$$

where

$$\mathcal{R}(\varphi) := \langle \widehat{e}_\varphi, [\![\Psi]\!] \rangle_{\mathcal{E}_h},$$

and

$$\mathcal{S}(\varphi) := \sum_{j=1}^{\mathcal{N}} \left[\widehat{e}_\varphi(x_j) - e_\varphi(x_j^-)\right] \xi_\Psi^+(x_j^-).$$

Proof. We only prove (2.37a) since the proofs of the other identities are similar. We begin by using the trivial identity

$$(e_w, \psi)_{\Omega_h} = (e_w, \Psi')_{\Omega_h} = (e_w, (\xi_\Psi^+)')_{\Omega_h} + (e_w, (\Pi^+\Psi)')_{\Omega_h}.$$

Next, we obtain an expression for $(e_w, (\Pi^+\Psi)')_{\Omega_h}$. Taking $v_1 = \Pi^+\Psi$ in the first error equation (2.23a), we get

$$\begin{aligned}(e_w, \psi)_{\Omega_h} &= (e_w, (\xi_\Psi^+)')_{\Omega_h} + \langle \widehat{e}_w, [\![\Pi^+\Psi]\!]\rangle_{\mathcal{E}_h} + (e_\theta + \kappa e_u - d^2 e_T, \Pi^+\Psi)_{\Omega_h} \\ &= -((\xi_w^+)', \xi_\Psi^+)_{\Omega_h} + (e_\theta + \kappa e_u - d^2 e_T, \Pi^+\Psi)_{\Omega_h} + \mathcal{T}(w)\end{aligned}$$

where

$$\mathcal{T}(w) := (e_w, (\xi_\Psi^+)')_{\Omega_h} + \langle \widehat{e}_w, [\![\Pi^+\Psi]\!]\rangle_{\mathcal{E}_h} + ((\xi_w^+)', \xi_\Psi^+)_{\Omega_h}.$$

It remains to show that $\mathcal{T}(w) = \mathcal{R}(w) - S(w)$. Integrating by parts the first term of the right-hand side, we get

$$\begin{aligned}\mathcal{T}(w) &= -(e_w', \xi_\Psi^+)_{\Omega_h} + \langle 1, [\![e_w \xi_\Psi^+]\!]\rangle_{\mathcal{E}_h} + \langle \widehat{e}_w, [\![\Pi^+\Psi]\!]\rangle_{\mathcal{E}_h} + ((\xi_w^+)', \xi_\Psi^+)_{\Omega_h} \\ &= -((e_w - \xi_w^+)', \xi_\Psi^+)_{\Omega_h} + \langle 1, [\![e_w \xi_\Psi^+]\!]\rangle_{\mathcal{E}_h} + \langle \widehat{e}_w, [\![\Pi^+\Psi]\!]\rangle_{\mathcal{E}_h}.\end{aligned}$$

By the definition of Π^+, we have that

$$((e_w - \xi_w^+)', \xi_\Psi^+)_{\Omega_h} = ((\Pi^+ e_w)', \xi_\Psi^+)_{\Omega_h} = 0.$$

Thus

$$\begin{aligned}
\mathcal{T}(w) &= \langle 1, [\![e_w \xi_\Psi^+]\!] \rangle_{\mathcal{E}_h} + \langle \widehat{e}_w, [\![\Pi^+ \Psi]\!] \rangle_{\mathcal{E}_h} \\
&= \langle 1, [\![e_w \xi_\Psi^+]\!] \rangle_{\mathcal{E}_h} + \langle 1, [\![\widehat{e}_w (\Pi^+ \Psi - \Psi)]\!] \rangle_{\mathcal{E}_h} + \langle 1, [\![\widehat{e}_w \Psi]\!] \rangle_{\mathcal{E}_h} \\
&= \langle 1, [\![(e_w - \widehat{e}_w) \xi_\Psi^+]\!] \rangle_{\mathcal{E}_h} + \mathcal{R}(w) \\
&= -\sum_{j=1}^{\mathcal{N}} \left[\widehat{e}_w(x_j) - e_w(x_j^-) \right] \xi_\Psi^+(x_j^-) + \mathcal{R}(w) \qquad \text{since } \xi_\Psi^+(x_{j-1}^+) = 0 \\
&= \mathcal{R}(w) - \mathcal{S}(w).
\end{aligned}$$

This completes the proof of Lemma 2.14. □

Step 2: Estimate of $\mathcal{S}(\varphi)$. In this step, we prove an auxiliary estimate of the term $\mathcal{S}(\varphi)$ for $\varphi \in \{T, N, M, \theta, u, w\}$.

Lemma 2.15. *With the same notation as in Lemma 2.14 we have*

$$|\mathcal{S}(\varphi)| \leq C h^{1/2} |\boldsymbol{e}|_{\mathcal{A}_h} \|\psi\|_0 .$$

Proof. We only prove the estimate for $\varphi = w$ since the proofs of the other estimates are similar. By definition of the numerical traces (4.2) we have

$$\widehat{e}_w(x_j) = (\{\!\{ e_w \}\!\} + C_{11} [\![e_w]\!] + C_{12} [\![e_u]\!] + C_{13} [\![e_\theta]\!] + C_{14} [\![e_M]\!] + C_{15} [\![e_N]\!] + C_{16} [\![e_T]\!])(x_j)$$

for any interior node x_j. Since $e_w(x_j^-) = \{\!\{ e_w \}\!\}(x_j) + [\![e_w]\!](x_j)/2$, we have for any interior node that

$$\begin{aligned}
\widehat{e}_w(x_j) - e_w(x_j^-) = ((C_{11} - 1/2) [\![e_w]\!] &+ C_{12} [\![e_u]\!] + C_{13} [\![e_\theta]\!] \\
&+ C_{14} [\![e_M]\!] + C_{15} [\![e_N]\!] + C_{16} [\![e_T]\!])(x_j).
\end{aligned}$$

Therefore,

$$\begin{aligned}\mathcal{S}(w) = \sum_{j=1}^{\mathcal{N}-1}((C_{11} \ &- 1/2)[\![e_w]\!] + C_{12}[\![e_u]\!] + C_{13}[\![e_\theta]\!] \\ &+ C_{14}[\![e_M]\!] + C_{15}[\![e_N]\!] + C_{16}[\![e_T]\!])(x_j)\xi_\Psi^+(x_j^-) \\ &- [\![e_w]\!](1)\,\xi_\Psi^+(1^-) \\ &:= \mathcal{S}_1 + \mathcal{S}_2 + \mathcal{S}_3 + \mathcal{S}_4 + S_5 + \mathcal{S}_6 + S_7,\end{aligned}$$

where

$$\begin{aligned}\mathcal{S}_1 &= \sum_{j=1}^{\mathcal{N}-1}(C_{11} - 1/2)[\![e_w]\!](x_j)\xi_\Psi^+(x_j^-), & \mathcal{S}_2 &= \sum_{j=1}^{\mathcal{N}-1} C_{12}[\![e_u]\!](x_j)\xi_\Psi^+(x_j^-), \\ \mathcal{S}_3 &= \sum_{j=1}^{\mathcal{N}-1} C_{13}[\![e_\theta]\!](x_j)\xi_\Psi^+(x_j^-), & \mathcal{S}_4 &= \sum_{j=1}^{\mathcal{N}-1} C_{14}[\![e_M]\!](x_j)\xi_\Psi^+(x_j^-), \\ \mathcal{S}_5 &= \sum_{j=1}^{\mathcal{N}-1} C_{15}[\![e_N]\!](x_j)\xi_\Psi^+(x_j^-), & \mathcal{S}_6 &= \sum_{j=1}^{\mathcal{N}-1} C_{16}[\![e_T]\!](x_j)\xi_\Psi^+(x_j^-),\end{aligned}$$

$$\mathcal{S}_7 = -[\![e_w]\!](1)\xi_\Psi^+(1^-).$$

Using assumption (3.4) and the approximation properties of Π^+ given in Lemma 2.13, the term $\mathcal{S}_1$ can be estimated as follows

$$\begin{aligned}|\mathcal{S}_1| &\le \Big(\sum_{j=1}^{\mathcal{N}-1}(C_{11} - 1/2)^2[\![e_w]\!]^2(x_j)\Big)^{1/2}\Big(\sum_{j=1}^{\mathcal{N}-1}(\xi_\Psi^+)^2(x_j^-)\Big)^{1/2} \\ &\le \Big(\sum_{j=1}^{\mathcal{N}-1}\mathsf{c}[\![e_w]\!]^2(x_j)\Big)^{1/2} Ch^{1/2}|\Psi|_{1,\Omega_h} \\ &\le Ch^{1/2}|\boldsymbol{e}|_{\mathcal{A}_h}\,\|\psi\|_0\,.\end{aligned}$$

In the last step we used the fact that $C_{61} = -\mathsf{c}$ and that $\Psi' = \psi$.

Similarly, the assumptions C_{12}^2, C_{13}^2, C_{14}^2, $C_{15}^2 \le \mathsf{c}$, and $C_{16} = -\mathsf{c}$, yield

$$|\mathcal{S}_i| \le Ch^{1/2}|\boldsymbol{e}|_{\mathcal{A}_h}\,\|\psi\|_0$$

for $i = 2, \ldots, 6$.

The estimate of $\mathcal{S}_7$ is as follows

$$\begin{aligned}|\mathcal{S}_7| &= \sqrt{\mathsf{c}}\,|[\![e_w]\!](1)| \cdot \frac{1}{\sqrt{\mathsf{c}}}|\xi_\Psi^+(1^-)| \\ &\le \left(\mathsf{c}[\![e_w]\!]^2(1)\right)^{1/2} Ch^{1/2}|\Psi|_1 \\ &\le Ch^{1/2}|\boldsymbol{e}|_{\mathcal{A}_h}\,\|\psi\|_0\,.\end{aligned}$$

This completes the proof. □

Step 3: Estimate of $\mathcal{R}(\varphi)$.

Lemma 2.16. *With the same notation as in Lemma 2.14 we have*

$$|\mathcal{R}(\varphi)| \le Ch^{k+1/2}\,\|\boldsymbol{e}\|_0\,|\boldsymbol{G}|_{k+1}\,\|\psi\|_0\,.$$

Proof. Let $x_j \in \mathcal{E}_h^\circ$ be an interior node. By the definition of Ψ, on element I_j, $\Psi(x) = \widetilde{\Psi}(x) - \widetilde{\Psi}(x_{j-1})$, and on element I_{j+1}, $\Psi(x) = \widetilde{\Psi}(x) - \widetilde{\Psi}(x_j)$. Thus, $\Psi(x_j^-) = \widetilde{\Psi}(x_j) - \widetilde{\Psi}(x_{j-1})$ and $\Psi(x_j^+) = \widetilde{\Psi}(x_j) - \widetilde{\Psi}(x_j)$ by the continuity of $\widetilde{\Psi}$. Thus,

$$[\![\Psi]\!](x_j) = \widetilde{\Psi}(x_j) - \widetilde{\Psi}(x_{j-1}) = (1,\psi)_{I_j} \le \sqrt{h_j}\,\|\psi\|_{0,I_j}\,. \tag{2.38}$$

For the boundary nodes,

$$[\![\Psi]\!](x_0) = -\Psi(x_0^+) = -[\widetilde{\Psi}(x_0) - \widetilde{\Psi}(x_0)] = 0, \tag{2.39}$$

and

$$[\![\Psi]\!](x_{\mathcal{N}}) = \Psi(x_{\mathcal{N}}^-) = \widetilde{\Psi}(x_{\mathcal{N}}) - \widetilde{\Psi}(x_{\mathcal{N}-1})] = (1,\psi)_{I_{\mathcal{N}}} \le \sqrt{h_{\mathcal{N}}}\,\|\psi\|_{0,I_{\mathcal{N}}}\,. \tag{2.40}$$

Now, by Lemma 2.6 and (2.38)-(2.40) we have

$$\begin{aligned}
|\mathcal{R}(\psi)| &= |\langle \widehat{e}_\varphi, [\![\Psi]\!] \rangle_{\mathcal{E}_h}| \\
&\leq \Big(\sum_{j=1}^{\mathcal{N}} (\widehat{e}_\varphi(x_j))^2 \Big)^{1/2} \Big(\sum_{j=1}^{\mathcal{N}} [\![\Psi]\!]^2(x_j) \Big)^{1/2} \\
&\leq C h^k \, \|\boldsymbol{e}\|_0 \, |\boldsymbol{G}|_{k+1} \Big(\sum_{j=1}^{\mathcal{N}} h_j \, \|\psi\|_{0,I_j}^2 \Big)^{1/2} \\
&\leq C h^{k+1/2} \, \|\boldsymbol{e}\|_0 \, |\boldsymbol{G}|_{k+1} \, \|\psi\|_0 \, .
\end{aligned}$$

This completes the proof □

Step 4: Estimate of $\Pi^+\Psi$.

Lemma 2.17. *With the same notation as in Lemma 2.14 we have*

$$\big\|\Pi^+\Psi\big\|_0 \leq C h \, \|\psi\|_0 \, .$$

Proof. Since, $\Pi^+\Psi = \Psi - \xi_\Psi^+$, we only need an estimate of $\|\Psi\|_0$. On an arbitrary element $I_j \in \Omega_h$ we have, by the definition of Ψ, that

$$\begin{aligned}
\|\Psi\|_{0,I_j} &= \big\|\Psi - \Psi(x_{j-1}^+)\big\|_{0,I_j} \quad \text{since } \Psi(x_{j-1}^+) = 0, \\
&= \big\|\Psi - \Pi_0^+\Psi\big\|_{0,I_j} \\
&= C h_j |\Psi|_{1,I_j}.
\end{aligned}$$

Here, Π_0^+ is the projection operator Π^+ with $k = 0$, and in the last step we made use of the approximation properties of Π_0^+ given in Lemma 2.13. Thus, adding over all elements, we get

$$\|\Psi\|_0 \leq C h |\Psi|_1 = C h \, \|\Psi'\|_0 = C h \, \|\psi\|_0 \, .$$

This completes the proof. □

Step 5: Estimate of $\|e_\varphi\|_0$.

Lemma 2.18. *We have, for any $\varphi \in \{T, N, M, \theta, u, w\}$, that*

$$\|e_\varphi\|_0^2 \le Ch^{k+1}|\boldsymbol{\varphi}|_{k+1} \|\boldsymbol{e}\|_0 + C(h + h^{k+1/2}|\boldsymbol{G}|_{k+1}) \|\boldsymbol{e}\|_0^2 .$$

Proof. We only show the details of how to estimate $\|e_w\|_0$, the proofs for the remaining variables follow similar lines. Taking $\psi = e_w$ in the representation formula (2.37a) we get

$$\|e_w\|_0^2 = -((\xi_w^+)', \xi_\Psi^+)_{\Omega_h} + (\mathcal{R} - \mathcal{S})(w) + (e_\theta + \kappa e_u - d^2 e_T, \Pi^+ \Psi)_{\Omega_h} \tag{2.41}$$

with the notation used in Lemma 2.14. By the approximation properties of Π^+ we have

$$\left\|(\xi_w^+)'\right\|_0 \le Ch^k |w|_{k+1} \le Ch^k |\boldsymbol{\varphi}|_{k+1},$$

and

$$\left\|\xi_\Psi^+\right\|_0 \le Ch|\Psi|_1 = Ch \|e_w\|_0 \le Ch \|\boldsymbol{e}\|_0$$

since $\Psi' = \psi = e_w$. Thus, by Cauchy-Schwarz inequality,

$$|((\xi_w^+)', \xi_\Psi^+)_{\Omega_h}| \le Ch^{k+1} |\boldsymbol{\varphi}|_{k+1} \|\boldsymbol{e}\|_0 . \tag{2.42}$$

By Lemma 2.15

$$|\mathcal{S}(w)| \le Ch^{1/2} |\boldsymbol{e}|_{\mathcal{A}_h} \|e_w\|_0 \le Ch^{1/2} |\boldsymbol{e}|_{\mathcal{A}_h} \|\boldsymbol{e}\|_0 ,$$

and hence by Lemma 2.5 we have

$$|\mathcal{S}(w)| \le Ch^{k+1} \|\boldsymbol{e}\|_0 |\boldsymbol{\varphi}|_{k+1} + Ch \|\boldsymbol{e}\|_0^2 . \tag{2.43}$$

By Lemma 2.16

$$|\mathcal{R}(w)| \le Ch^{k+1/2} \|\boldsymbol{e}\|_0 |\boldsymbol{G}|_{k+1} \|e_w\| \le Ch^{k+1/2} \|\boldsymbol{e}\|_0^2 |\boldsymbol{G}|_{k+1}. \tag{2.44}$$

By Lemma 2.17 and Cauchy-Schwarz inequality we have

$$\begin{aligned} |(e_\theta + \kappa e_u - d^2 e_T, \Pi^+ \Psi)_{\Omega_h}| &\leq (\|e_\theta\|_0 + \|\kappa e_u\|_0 + \left\|d^2 e_T\right\|_0) \cdot Ch \|e_w\|_0 \\ &\leq Ch \|\boldsymbol{e}\|_0^2 , \end{aligned} \tag{2.45}$$

where we have used the boundedness of κ, the fact that $d < 1$, and that $\|e_\varphi\|_0 \leq \|\boldsymbol{e}\|_0$ for any $\varphi \in \{T, \theta, u, w\}$.

Inserting the estimates (2.42)–(2.45) into (2.41) yields the desired estimate. □

2.3.6 Proof of Theorem 2.3

Applying the estimate in Lemma 2.18 for all $\varphi \in \{T, N, M, \theta, u, w\}$, and adding the resulting estimates we get that

$$\|\boldsymbol{e}\|_0^2 \leq Ch^{k+1} |\boldsymbol{\varphi}|_{k+1} \|\boldsymbol{e}\|_0 + C(h + h^{k+1/2} |\boldsymbol{G}|_{k+1}) \|\boldsymbol{e}\|_0^2 .$$

Assuming h is small enough so that

$$C(h + h^{k+1/2} |\boldsymbol{G}|_{k+1}) \leq \alpha < 1$$

for some constant α, we see that

$$\|\boldsymbol{e}\|_0^2 \leq Ch^{k+1} |\boldsymbol{\varphi}|_{k+1} \|\boldsymbol{e}\|_0 .$$

Canceling $\|\boldsymbol{e}\|_0$ on both sides yields the estimate (2.21).

The error estimate in the energy seminorm, namely, (2.20) now follows from inserting (2.21) into the estimate in Lemma 2.5.

This finishes the proof of Theorem 2.3.

2.3.7 Proof of Theorem 3.1

This is a simple implication of Lemma 2.6 and Theorem 2.3.

2.4 Numerical Results

In this section, we display numerical results verifying our theoretical findings. We solve the equations (2.2)–(2.5) with $\kappa \equiv 1$, together with boundary conditions $w = u = \theta = 0$ at $\partial\Omega$. Although, the theory has been carried out for variable curvature, we take a constant κ so that we can compute the exact solution and produce history of convergence tables. This choice corresponds to a circular arch of thickness d. To verify that the DG method is locking-free, d is taken to be 10^{-1}, 10^{-4}, and finally decreased down to 10^{-8}. Observe that, since this parameter only appears as d^2 in the model, from a computational perspective the last choice is equivalent to the limiting case in which we consider an arch of *thickness zero*. We take uniform loading in arc length, namely, $p = q = 1$ in Ω.

The DG method is defined by the weak formulation (2.2) whose numerical traces are given by the formulas (4.2)–(2.5) which are obtained by setting

$$C_{16} = C_{25} = C_{34} = C_{43} = C_{52} = C_{61} = -1$$

for all x in $\mathcal{E}_h$, except $C_{16} = C_{25} = C_{34} = 0$ on $\partial\Omega$, and setting all the other coefficients to zero.

We display our results in Tables 1 through 3. Therein k indicates the polynomial degree we used to define the DG method, and "mesh = i" means we employed a uniform mesh with 2^i elements. We also display the numerical orders of convergence which are computed as follows. Let $\boldsymbol{e}(i)$ denote the error where a mesh with 2^i elements have been used to obtain the DG solution. The approximate order of convergence, r_i, at the level i is defined as $r_i = \log(\boldsymbol{e}(i-1)/\boldsymbol{e}(i))/\log 2$.

We see that the optimal rates of convergence in L^2-norm and the $k+1/2$-order convergence

Table 1: History of convergence in the energy seminorm.

		$d = 10^{-1}$		$d = 10^{-4}$		$d = 10^{-8}$	
k	mesh	$\lvert\boldsymbol{e}\rvert_{\mathcal{A}_h}$	order	$\lvert\boldsymbol{e}\rvert_{\mathcal{A}_h}$	order	$\lvert\boldsymbol{e}\rvert_{\mathcal{A}_h}$	order
	7	1.04E-01	0.39	8.63E-02	0.44	8.63E-02	0.44
0	8	7.78E-02	0.42	6.25E-02	0.47	6.25E-02	0.47
	9	5.70E-02	0.45	4.47E-02	0.48	4.47E-02	0.48
	6	2.58E-04	1.49	1.73E-04	1.49	1.73E-04	1.49
1	7	9.17E-05	1.50	6.12E-05	1.50	6.12E-05	1.50
	8	3.25E-05	1.50	2.17E-05	1.50	2.17E-05	1.50
	6	9.43E-07	2.47	7.24E-07	2.47	7.24E-07	2.47
2	7	1.68E-07	2.49	1.29E-07	2.49	1.29E-07	2.49
	8	2.99E-08	2.49	2.29E-08	2.49	2.29E-08	2.49
	5	6.64E-09	3.49	4.72E-09	3.49	4.72E-09	3.49
3	6	5.90E-10	3.49	4.18E-10	3.49	4.18E-10	3.49
	7	5.22E-11	3.50	3.71E-11	3.50	3.71E-11	3.50

Table 2: History of convergence in the L^2-norm.

		$d = 10^{-1}$		$d = 10^{-4}$		$d = 10^{-8}$	
k	mesh	$\|\boldsymbol{e}\|_0$	order	$\|\boldsymbol{e}\|_0$	order	$\|\boldsymbol{e}\|_0$	order
	8	2.32E-01	0.66	4.77E-02	0.88	4.77E-02	0.88
0	9	1.34E-01	0.79	2.49E-02	0.94	2.49E-02	0.94
	10	7.27E-02	0.88	1.27E-02	0.97	1.27E-02	0.97
	7	2.41E-06	2.09	1.58E-06	2.02	1.58E-06	2.02
1	8	5.92E-07	2.03	3.94E-07	2.01	3.94E-07	2.01
	9	1.47E-07	2.01	9.81E-08	2.00	9.81E-08	2.00
	6	5.31E-08	2.98	4.08E-08	2.98	4.08E-08	2.98
2	7	6.67E-09	2.99	5.13E-09	2.99	5.13E-09	2.99
	8	8.37E-10	3.00	6.43E-10	2.99	6.44E-10	2.99
	5	4.58E-10	4.01	3.25E-10	4.01	3.25E-10	4.01
3	6	2.85E-11	4.01	2.02E-11	4.01	2.02E-11	4.01
	7	1.78E-12	4.00	1.26E-12	4.00	1.26E-12	4.00

Table 3: History of convergence of the numerical traces.

		$d = 10^{-1}$		$d = 10^{-4}$		$d = 10^{-8}$	
k	mesh	$\|\widehat{e}\|_{L^\infty}$	order	$\|\widehat{e}\|_{L^\infty}$	order	$\|\widehat{e}\|_{L^\infty}$	order
	8	2.32E-01	0.66	4.55E-02	0.88	4.55E-02	0.88
0	9	1.34E-01	0.79	2.38E-02	0.94	2.38E-02	0.94
	10	7.25E-02	0.88	1.22E-02	0.97	1.22E-02	0.97
	6	3.70E-06	3.02	3.83E-07	3.03	3.83E-07	3.03
1	7	4.60E-07	3.01	4.74E-08	3.02	4.74E-08	3.02
	8	5.72E-08	3.01	5.89E-09	3.01	5.89E-09	3.01
	6	5.35E-12	4.94	6.04E-12	4.96	6.04E-12	4.96
2	7	1.71E-13	4.97	1.91E-13	4.98	1.91E-13	4.98
	8	5.40E-15	4.98	6.02E-15	4.99	6.02E-15	4.99
	6	4.79E-18	7.06	4.98E-20	7.72	4.98E-20	7.72
3	7	3.66E-20	7.03	2.62E-22	7.57	2.62E-22	7.57
	8	2.82E-22	7.02	1.55E-24	7.40	1.55E-24	7.40

in the energy seminorm predicted by Theorem 2.3 ore indeed achieved. The results in Table 1 also shows that the estimate 2.20 is actually sharp. We also see from Table 3 that all the numerical traces superconverge with order $2k+1$ at the nodes of the mesh, in perfect agreement with Theorem 3.1.

As predicted by our error estimates in Section 4.5 the DG method is completely robust with respect to the thickness of the arch, and the method is free from locking.

Next, we display an example where we compute the DG solution where the curvature of the arch, $\kappa = \kappa(x)$, is variable. Since computing the closed form of the exact solution to (1.3) is impossible for the parabolic arch we will describe, we only display a plot of the undeformed configuration of the arch, and its deformed configuration after the application of the loads The undeformed arch is given by the formula $y(x) = 1 - x^2$ for $x \in [-1, 1]$. Its arc length parametrization is then

$$s(x) = \int_{-1}^{x} \sqrt{1 + 4t^2}\, dt.$$

The curvature at x is $\kappa(x) = -2/(1+4x^2)^{3/2}$. Thus, the curvature at the arch length s is

$$\kappa(s) = -\frac{2}{(1+4x^2)^{3/2}}, \qquad s \in [0, L].$$

Here, $L = s(1)$ is the total arc length of the arch. We consider an arch of thickness $d = 0.1$. The tangential and transverse loads are taken, respectively, as

$$p(s) = \frac{4x(s)}{1+4x^2(s)}, \qquad q(s) = \frac{-10}{1+4x^2(s)}.$$

A simple computation shows that these correspond to a slight scaling of an arch loaded uniformly in horizontal direction. We see from Figure 3 that the total displacement produced by the DG approximation seems to be reasonable.

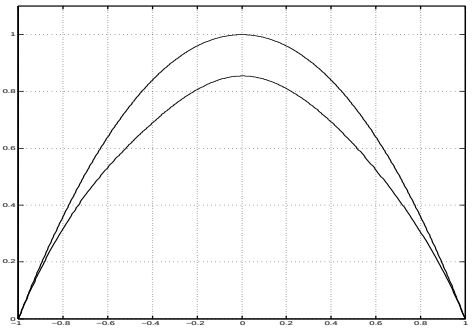

Figure 3: A parabolic arch. Undeformed configuration (top curve) and the DG solution after the application of the loads (bottom curve).

2.5 Concluding Remarks

We have devised a general family of DG methods for a Naghdi type arch model, and provided conditions under which the DG approximation is well defined. We then restricted ourselves to a particular subfamily of DG methods, and proved that the approximate solution converges optimally for all the unknowns. We have also shown that these methods are free from shear locking since the error estimates are independent of the thickness of the arch. A superconvergence property of the numerical traces was also proved. All of these results can be considered as extensions of those for DG methods for Timoshenko beams studied in [20] and [15].

A rightful criticism for the methods studied in this paper is the proliferation of the number of degrees of freedom involved in the DG formulation. Such a criticism can be removed by considering the so-called hybridizable DG (HDG) methods which allows the elimination of the internal degrees of freedom from the final linear system and obtaining an equivalent formulation only in terms of the nodal degrees of freedom for only three of the unknowns.

Celiker *et al.* carried out the details of such a simplification in the context of DG methods for Timoshenko beams, see [19] and [18]. Therein they have devised and analyzed a wide class of HDG methods for Timoshenko beams and they showed that they are optimally convergent and are free from shear locking.

3 Element-by-element post-processing

3.1 Introduction

In chapter 2, a family of locking-free discontinuous Galerkin (DG) methods for a Naghdi-type arch model was introduced. They have proved that the approximation converges with order $k+1$ when polynomials of degree k are used. In this section, we construct an element-by-element post-processing that converges remarkably faster.

This post-processing is based on the fact that a superconvergence phenomenon takes place at the nodes of the mesh. Indeed, the numerical traces of the DG method converge to the nodal values of the exact solution with order $2k+1$ when polynomials of degree k are used to compute the DG approximation, see [42]. The main goal of this paper is to exploit this phenomenon to post-process the DG solution element-by-element and obtain a better solution which superconverges to the exact solution with order $2k+1$ in the L^2-norm throughout the domain rather than at merely some isolated points of the mesh.

A similar superconvergent post-processing result has been proved for DG methods for convection-diffusion problems in [51]. Based on the superconvergence result proved therein, Cockburn and Ichikawa [52] devised a post-processing for the approximation of linear functionals which is superconvergent of order $4k+1$. In [16] Celiker and Cockburn designed a post-processing for DG methods for Timoshenko beams which is superconvergent of order $2k+1$ in the L^∞-norm throughout the computational domain. This result was based on the numerical observation that the numerical traces of the DG approximation for Timoshenko beams are also superconvergent of order $2k+1$ at the nodes of the mesh. Shortly later, the superconvergence of the numerical traces was put on a firm mathematical ground in [20].

As we will describe below, the Timoshenko beam model can be viewed as a special case of the Naghdi arch model where the beam is considered as an arch with zero curvature. The post-processing we display in this paper is thus inspired by the one introduced in [16]. Despite this close similarity, the coupling of some of the unknowns in the Naghdi arch model renders both the post-processing and its error analysis more involved. This is especially the case for the latter because it requires the analysis of a linear system of initial value problems whose solution is approximated by using approximate data. This is the main reason why we prove an L^2-error estimate for the post-processed approximation unlike the L^∞-error estimate for the Timoshenko beam post-processing. Notwithstanding, it is possible to prove an L^∞-error estimate at the expense of requiring high order regularity, following, for example, [53, 54].

3.2 Post-processing

Next, we describe the post-processing

$$\boldsymbol{\varphi}_h^* := (T_h^*, N_h^*, M_h^*, \theta_h^*, u_h^*, w_h^*)$$

of the approximate solution $\boldsymbol{\varphi}_h = (T_h, N_h, M_h, \theta_h, u_h, w_h)$ provided by the DG method. It is based on the fact that the numerical traces superconverge at each of the nodes with order $2k+1$. To state this result we need to introduce some notation. We define the error of approximation as

$$e_\varphi = \varphi - \varphi_h, \qquad \widehat{e}_\varphi = \varphi - \widehat{\varphi}_h,$$

for any $\varphi \in \{T, N, M, \theta, u, w\}$, and set

$$\boldsymbol{e} = \boldsymbol{\varphi} - \boldsymbol{\varphi}_h, \qquad \widehat{\boldsymbol{e}} = \boldsymbol{\varphi} - \widehat{\boldsymbol{\varphi}}_h.$$

Here

$$\boldsymbol{\varphi} := (T, N, M, \theta, u, w)$$

denotes the exact solution of the governing equations (1.3). The error in the numerical traces of φ_h is defined as

$$\|\widehat{e}_\varphi\|_\infty := \|\widehat{e}_\varphi\|_{\ell^\infty(\mathcal{E}_h)} := \max_{x_j \in \mathcal{E}_h} |\widehat{e}_\varphi(x_j)|,$$

and the global error in the numerical traces is set to be

$$\|\widehat{\boldsymbol{e}}\|_\infty := \max_{\varphi \in \{T,N,M,\theta,u,w\}} \|\widehat{e}_\varphi\|_\infty .$$

We denote by $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$ the usual norm and seminorm, respectively, in the Sobolev space $H^s(D)$ where D is any subset of Ω_h. We drop the subindex D whenever $D = \Omega_h$ or $D = \Omega$. We set, for $\boldsymbol{u} = (u_1, u_2, u_3, u_4, u_5, u_6)$,

$$|\boldsymbol{u}|_{s,D} := (|u_1|^2_{s,D} + |u_2|^2_{s,D} + |u_3|^2_{s,D} + |u_4|^2_{s,D} + |u_5|^2_{s,D} + |u_6|^2_{s,D})^{1/2}.$$

In [42] the following wide family of DG methods has been analyzed. They are defined by setting the functions C_{ij} as follows.

$$C_{16} = C_{25} = C_{34} = C_{43} = C_{52} = C_{61} = -\mathsf{c} \tag{3.1}$$

for all x in $\mathcal{E}_h$, except

$$C_{16} = C_{25} = C_{34} = 0 \quad \text{on} \quad \partial\Omega. \tag{3.2}$$

Here, $\mathsf{c} > 0$ is any constant which is independent of the mesh size h. We assume that

$$C_{ij}^2 \leq \mathsf{c} \quad \text{for all } i, j = 1, \ldots, 6, \tag{3.3}$$

and that

$$(C_{ii}(x) - 1/2)^2 \leq \mathsf{c} \quad \text{for all } i = 1, \ldots, 6. \tag{3.4}$$

Such a choice can be obtained, for example, by setting

$$C_{16} = C_{25} = C_{34} = C_{43} = C_{52} = C_{61} = -1$$

for all x in $\mathcal{E}_h$, except

$$C_{16} = C_{25} = C_{34} = 0 \quad \text{on} \quad \partial\Omega,$$

and setting all the remaining C_{ij}'s to zero.

We are now ready to state the superconvergence result for the numerical traces.

Theorem 3.1. ([42]) *Let $k \geq 0$ be a polynomial degree and suppose that $\boldsymbol{\varphi}$ belongs to $[H^{k+1}(\Omega_h)]^6$. Let $\boldsymbol{\varphi}_h$ be the DG solution defined by the weak formulation* (2.2)*, and the numerical traces* (4.2)–(2.5) *where the functions C_{ij} are defined so as to satisfy* (3.1)–(3.4)*. Then,*

$$\|\boldsymbol{\varphi} - \widehat{\boldsymbol{\varphi}}_h\|_\infty \leq C\ h^{2k+1} |\boldsymbol{\varphi}|_{k+1} \tag{3.5}$$

for some constant C independent of h and d.

Our post-processing is defined in an element-by-element fashion as follows. On the element $I_j = (x_{j-1}, x_j)$, $1 \leq j \leq \mathcal{N}$, we define the post-processed solution

$$\boldsymbol{\varphi}_h^* = (T_h^*, N_h^*, M_h^*, \theta_h^*, u_h^*, w_h^*)$$

as the element of the space $[\mathcal{P}^{2k}(I_j)]^6$ in four simple steps as follows.

Step 1: Compute T_h^* and N_h^* by solving

$$-(T_h^*, v_1')_{I_j} + T_h^*(x_j^-)v_1(x_j^-) + (\kappa N_h^*, v_1)_{I_j} = (q, v_1)_{I_j} + \widehat{T}_h(x_{j-1})v_1(x_{j-1}^+), \tag{3.6a}$$

$$-(N_h^*, v_2')_{I_j} + N_h^*(x_j^-)v_2(x_j^-) - (\kappa T_h^*, v_2)_{I_j} = (p, v_2)_{I_j} + \widehat{N}_h(x_{j-1})v_2(x_{j-1}^+), \tag{3.6b}$$

for all v_1 and v_2 in $\mathcal{P}^{2k}(I_j)$.

Step 2: Compute M_h^* by solving

$$-(M_h^*, v_3')_{I_j} + M_h^*(x_j^-)v_3(x_j^-) = (T_h^*, v_3)_{I_j} + \widehat{M}_h(x_{j-1})v_3(x_{j-1}^+), \tag{3.7}$$

for all v_3 in $\mathcal{P}^{2k}(I_j)$.

Step 3: Compute θ_h^* by solving

$$-(\theta_h^*, v_4')_{I_j} + \theta_h^*(x_j^-)v_4(x_j^-) = (M_h^*, v_4)_{I_j} + \widehat{\theta}_h(x_{j-1})v_4(x_{j-1}^+), \tag{3.8}$$

for all v_4 in $\mathcal{P}^{2k}(I_j)$.

Step 4: Compute u_h^* and w_h^* by solving

$$\begin{aligned} -(u_h^*, v_5')_{I_j} + u_h^*(x_j^-)v_5(x_j^-) - (\kappa w_h^*, v_5)_{I_j} \\ = d^2(N_h^*, v_5)_{I_j} + \widehat{u}_h(x_{j-1})v_5(x_{j-1}^+), \end{aligned} \tag{3.9a}$$

$$\begin{aligned} -(w_h^*, v_6')_{I_j} + w_h^*(x_j^-)v_6(x_j^-) + (\kappa u_h^*, v_6)_{I_j} \\ = d^2(T_h^*, v_6)_{I_j} - (\theta_h^*, v_6)_{I_j} + \widehat{w}_h(x_{j-1})v_6(x_{j-1}^+), \end{aligned} \tag{3.9b}$$

for all v_5 and v_6 in $\mathcal{P}^{2k}(I_j)$.

Next, we state a theorem about the existence and uniqueness of the post-processed solution.

Theorem 3.2. *Consider the post-processing defined by* (3.6)–(3.9) *on an arbitrary element* $I_j \in \Omega_h$. *These equations define a unique solution* $\boldsymbol{\varphi}_h^* = (T_h^*, N_h^*, M_h^*, \theta_h^*, u_h^*, w_h^*)$ *provided that the condition* (2.10) *is satisfied whenever* κ *is not identically equal to a constant on* I_j.

Remark. If κ is identically constant, i.e. the arch is locally circular or flat, on an element I_j then the condition (2.10) is not necessary, and the post-processing automatically defines a unique solution.

It is not difficult to see that the equations (3.6)–(3.9) are the discretization by the classical DG method [55, 56] of the following system of initial value problems

$$(T^*)' + \kappa N^* = q \qquad \text{in } I_j, \qquad T^*(x_{j-1}) = \widehat{T}_h(x_{j-1}), \tag{3.10a}$$

$$(N^*)' - \kappa T^* = p \qquad \text{in } I_j, \qquad N^*(x_{j-1}) = \widehat{N}_h(x_{j-1}), \tag{3.10b}$$

$$(M^*)' = T^* \qquad \text{in } I_j, \qquad M^*(x_{j-1}) = \widehat{M}_h(x_{j-1}), \tag{3.10c}$$

$$(\theta^*)' = M^* \qquad \text{in } I_j, \qquad \theta^*(x_{j-1}) = \widehat{\theta}_h(x_{j-1}), \tag{3.10d}$$

$$(u^*)' - \kappa w^* = d^2 N^* \qquad \text{in } I_j, \qquad u^*(x_{j-1}) = \widehat{u}_h(x_{j-1}), \tag{3.10e}$$

$$(w^*)' + \kappa u^* = d^2 T^* - \theta^* \qquad \text{in } I_j, \qquad w^*(x_{j-1}) = \widehat{w}_h(x_{j-1}). \tag{3.10f}$$

Its step-by-step nature reveals that when defining the post-processing (3.6)–(3.9) we made use of the fact that the system of equations (3.10) is partially decoupled in the following sense. It is possible to solve for T^* and N^* using only the equations (3.10a) and (3.10b). Then we can insert T^* into (3.10c) and solve for M^*, and then insert M^* into (3.10d) to solve for θ^*. Finally, we may insert N^* into (3.10e), and T^* and θ^* into (3.10f), and solve for u^* and w^*.

Based on the above observation, we can rewrite (3.10) in a single framework as follows:

$$(\boldsymbol{\varphi}_\ell^*)' - A_\ell \boldsymbol{\varphi}_\ell^* = \boldsymbol{f}_\ell^* \quad \text{in } I_j, \qquad \boldsymbol{\varphi}_\ell^*(x_{j-1}) = \widehat{\boldsymbol{\varphi}}_\ell(x_{j-1}) \tag{3.11}$$

for $\ell = 1, 2, 3, 4$. Here,

$$\boldsymbol{\varphi}_1^* := \begin{bmatrix} T^* \\ N^* \end{bmatrix}, \qquad \boldsymbol{\varphi}_2^* := [\, M^*], \quad \boldsymbol{\varphi}_3^* := [\, \theta^*], \quad \boldsymbol{\varphi}_4^* := \begin{bmatrix} u^* \\ w^* \end{bmatrix},$$

and similarly for $\widehat{\boldsymbol{\varphi}}^*_\ell$,

$$A_1 := \begin{bmatrix} 0 & -\kappa \\ \kappa & 0 \end{bmatrix}, \quad A_2 = [0], \qquad A_3 = [0], \qquad A_4 = \begin{bmatrix} 0 & \kappa \\ -\kappa & 0 \end{bmatrix},$$

$$\boldsymbol{f}^*_1 = \begin{bmatrix} q \\ p \end{bmatrix}, \qquad \boldsymbol{f}^*_2 := [T^*], \quad \boldsymbol{f}^*_3 := [M^*], \quad \boldsymbol{f}^*_4 := \begin{bmatrix} d^2N^* \\ d^2T^* - \theta^* \end{bmatrix}.$$

Consequently, we can reformulate the post-processing defined by the equations (3.6)–(3.9) in the following unified framework. Find $(\boldsymbol{\varphi}^*_{1,h}, \boldsymbol{\varphi}^*_{2,h}, \boldsymbol{\varphi}^*_{3,h}, \boldsymbol{\varphi}^*_{4,h},) \in [\mathcal{P}^{2k}(I_j)]^2 \times \mathcal{P}^{2k}(I_j) \times \mathcal{P}^{2k}(I_j) \times [\mathcal{P}^{2k}(I_j)]^2$ such that

$$\begin{aligned} -(\boldsymbol{\varphi}^*_{\ell,h}, \boldsymbol{v}'_\ell)_{I_j} + \boldsymbol{\varphi}^*_{\ell,h}(x_j^-) \cdot \boldsymbol{v}_\ell(x_j^-) - (A_\ell \boldsymbol{\varphi}^*_{\ell,h}, \boldsymbol{v}_\ell)_{I_j} \\ = (\boldsymbol{f}^*_\ell, \boldsymbol{v}_\ell)_{I_j} + \widehat{\boldsymbol{\varphi}}_{\ell,h}(x_{j-1}) \cdot \boldsymbol{v}_\ell(x_{j-1}^+) \end{aligned} \tag{3.12}$$

for all $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4) \in [\mathcal{P}^{2k}(I_j)]^2 \times \mathcal{P}^{2k}(I_j) \times \mathcal{P}^{2k}(I_j) \times [\mathcal{P}^{2k}(I_j)]^2$. Here we have used the obvious definitions of $\boldsymbol{\varphi}^*_{\ell,h}$ and $\widehat{\boldsymbol{\varphi}}_{\ell,h}$, and A_ℓ and $\boldsymbol{f}^*_\ell$ are the same as above. We have also employed the following notation. For two vector-valued functions $\boldsymbol{\varphi}$ and $\boldsymbol{v}$ in $[L^2(I_j)]^m$

$$(\boldsymbol{\varphi}, \boldsymbol{v})_{I_j} := \int_{I_j} \boldsymbol{\varphi} \cdot \boldsymbol{v} = \sum_{i=1}^m \int_{I_j} \varphi_i v_i,$$

and " $\cdot$ " denotes the usual dot product of two vectors in $\mathbb{R}^m$.

Next, we state our main result.

Theorem 3.3. *Under the hypotheses of Theorem 3.1, the error of the post-processed approximation is such that*

$$\|\boldsymbol{\varphi} - \boldsymbol{\varphi}^*_h\|_{0,\Omega_h} \leq C\, h^{2k+1} \tag{3.13}$$

for some constant C independent of h and d.

Remark. This theorem extends earlier results by Celiker and Cockburn for DG methods for convection-diffusion problems in [51], and for Timoshenko beams in [16]. The main difficulty here arises from considering an arbitrary geometry for the arch which results in the appearance of the additional variables u and N in the governing equations. Moreover, the transverse displacement u is coupled with the tangential displacement w, and the shear stress T is coupled with the membrane stress N, as can be seen from (1.3a)–(1.3b) and (1.3e)–(1.3f), respectively. Consequently, for the post-processing we have to solve a system of equations, rather than a set of scalar equations, as is evident from (3.11). This renders the analysis of the post-processing of DG methods for arches considerably more involved than that of the DG methods for beams. Let us note that extending a result for beams to one for arches is analogous to extending a result for plates to one for shells and hence poses several challenges.

Remark. Since the constant C appearing in the estimate (3.13) is independent of the thickness parameter d, the post-processed solution is free from shear and membrane locking.

Remark. The estimate (3.13) shows that the post-processed approximation converges with order $2k+1$ throughout the computational domain. This should be contrasted with the fact that before post-processing the approximation converges with the optimal order or $k+1$. Hence, for $k \geq 1$, the order of convergence is almost doubled by the local post-processing.

Remark. The value of the increase in the convergence order mentioned in the above remark becomes more evident if we calculate the computational cost of this post-processing. Since it is performed in an element-by-element fashion the total cost is $\mathcal{N}$ times the cost on one element. Therefore it is extremely inexpensive. More explicitly, Steps 1 and 4 require solving linear systems of order $2(2k+1)$, and Steps 2 and 3 can be performed by inverting a single

linear system of order $2k+1$. It is thus easy to see that the computational cost of the post-processing is negligible when compared to that of computing the original DG solution which, in general, requires solving a linear system of order $6\mathcal{N}(k+1)$.

3.3 Proofs

In this section we give detailed proofs of our results in Section 3.2. We begin with the proof of Theorem 3.2. It is based on the following lemma which was proved in [42]. We also provide a proof here for the sake of completeness.

Lemma 3.4. *Let r be a non-negative integer. Let $f, g \in \mathcal{P}^r([a,b])$ be such that*

$$f(a) = g(a) = 0. \tag{3.14}$$

Suppose that

$$\mathsf{P}_r(g' + \alpha f) = 0 \quad \text{and} \quad \mathsf{P}_r(f' - \alpha g) = 0, \tag{3.15}$$

where α is a function in $L^\infty([a,b])$ and P_r denotes the L^2-orthogonal projection into $\mathcal{P}^r([a,b])$. Then $f = g = 0$ in $[a,b]$ if

(a) α is identically equal to a constant, or

(b) α is not identically equal to a constant and

$$b - a \leq \frac{1}{2\left\|\alpha - \overline{\alpha}\right\|_{L^\infty([a,b])}} \tag{3.16}$$

where $\overline{\alpha}$ denotes the average value of α over the interval $[a,b]$.

Proof. By (A.3), we have that

$$g' + \mathsf{P}_r(\alpha f) = 0, \tag{3.17a}$$

$$f' - \mathsf{P}_r(\alpha g) = 0, \tag{3.17b}$$

pointwise on $[a, b]$. Multiplying (3.17a) by g and (3.17b) with f we get

$$\frac{1}{2}(g^2)' + g\mathsf{P}_r(\alpha f) = 0, \qquad \frac{1}{2}(f^2)' - f\mathsf{P}_r(\alpha g) = 0,$$

and hence

$$\frac{1}{2}(g^2 + f^2)' = f\mathsf{P}_r(\alpha g) - g\mathsf{P}_r(\alpha f) = f\mathsf{P}_r((\alpha - \overline{\alpha})g) - g\mathsf{P}_r((\alpha - \overline{\alpha})f) \tag{3.18}$$

since $-f\mathsf{P}_r(\overline{\alpha}g) + g\mathsf{P}_r(\overline{\alpha}f) = 0$ because $\overline{\alpha}$ is a constant and $f, g \in \mathcal{P}^r([a, b])$. Integrating both sides of (3.18) from a to an arbitrary x in $[a, b]$, and using (A.1), we obtain

$$\frac{1}{2}(g^2 + f^2)(x) = T_1(x) + T_2(x)$$

where

$$T_1(x) = \int_a^x f(s)\mathsf{P}_r((\alpha - \overline{\alpha})g)(s)\, ds, \qquad T_2(x) = -\int_a^x g(s)\mathsf{P}_r((\alpha - \overline{\alpha})f)(s)\, ds.$$

By Cauchy-Schwarz inequality

$$\begin{aligned} |T_1(x)| &\leq \|f\|_{L^2([a,b])} \|(\alpha - \overline{\alpha})g\|_{L^2([a,b])} \\ &\leq \|\alpha - \overline{\alpha}\|_{L^\infty([a,b])} \|f\|_{L^2([a,b])} \|g\|_{L^2([a,b])}. \end{aligned}$$

Similarly,

$$|T_2(x)| \leq \|\alpha - \overline{\alpha}\|_{L^\infty([a,b])} \|f\|_{L^2([a,b])} \|g\|_{L^2([a,b])},$$

and hence

$$\frac{1}{2}(g^2 + f^2)(x) \leq 2\|\alpha - \overline{\alpha}\|_{L^\infty([a,b])} \|f\|_{L^2([a,b])} \|g\|_{L^2([a,b])}.$$

Integrating both sides over $x \in [a,b]$ implies

$$\begin{aligned}\frac{1}{2}(\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])}) &\leq 2(b-a)\,\|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\,\|f\|_{L^2([a,b])}\,\|g\|_{L^2([a,b])} \\ &\leq (b-a)\,\|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\,(\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])})\end{aligned}$$

by Young's inequality. Thus,

$$\left[\frac{1}{2} - (b-a)\,\|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\right](\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])}) \leq 0. \tag{3.19}$$

Now, if α is identically constant on $[a,b]$ then $\overline{\alpha} = \alpha$ and the result follows since in such a case (A.7) implies $\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])} = 0$. If α is not identically constant on $[a,b]$ then we reach the same conclusion by (A.4).

This completes the proof. □

We are now ready to prove Theorem 3.2.

Proof. (Theorem 3.2) We only prove the existence and uniqueness of Step 1 of the post-processing. Steps 2 and 3 are well defined since they are nothing but the classical DG method applied to first order problems on a single element. Step 4 is almost identical to Step 1.

Due to the linearity of the problem it suffices to show that the only solution to (3.6) with

$$p = q = 0 \quad \text{in } I_j,$$

and

$$\widehat{T}_h(x_{j-1}) = \widehat{N}_h(x_{j-1}) = 0,$$

is

$$T_h^* = N_h^* = 0 \quad \text{in } I_j.$$

In this case, the equations (3.6) simplify to

$$-(T_h^*, v_1')_{I_j} + T_h^*(x_j^-)v_1(x_j^-) + (\kappa N_h^*, v_1)_{I_j} = 0, \tag{3.20a}$$

$$-(N_h^*, v_2')_{I_j} + N_h^*(x_j^-)v_2(x_j^-) - (\kappa T_h^*, v_2)_{I_j} = 0, \tag{3.20b}$$

Taking $v_1 = T_h^*$ in (3.20a) and $v_2 = N_h^*$ in (3.20b), and adding the resulting equations we get

$$-(T_h^*, (T_h^*)')_{I_j} + (T_h^*(x_j^-))^2 - (N_h^*, (N_h^*)')_{I_j} + (N_h^*(x_j^-))^2 = 0.$$

This implies,

$$\frac{1}{2}\left[(T_h^*)^2(x_{j-1}^+) + (T_h^*)^2(x_j^-)\right] + \frac{1}{2}\left[(N_h^*)^2(x_{j-1}^+) + (N_h^*)^2(x_j^-)\right] = 0.$$

Hence,

$$T_h^*(x_{j-1}^+) = T_h^*(x_j^-) = N_h^*(x_{j-1}^+) = N_h^*(x_j^-) = 0. \tag{3.21}$$

This further simplifies (3.20) to

$$-(T_h^*, v_1')_{I_j} + (\kappa N_h^*, v_1)_{I_j} = 0,$$

$$-(N_h^*, v_2')_{I_j} - (\kappa T_h^*, v_2)_{I_j} = 0,$$

Upon a simple integration by parts and invoking (3.21) we get that

$$((T_h^*)' + \kappa N_h^*, v_1)_{I_j} = 0, \quad \text{and} \quad ((N_h^*)' - \kappa T_h^*, v_2)_{I_j} = 0.$$

for all v_1 and v_2 in $\mathcal{P}^r([a,b])$. In other words,

$$\mathsf{P}_r((T_h^*)' + \kappa N_h^*, v_1) = 0, \quad \text{and} \quad \mathsf{P}_r((N_h^*)' - \kappa T_h^*, v_2) = 0.$$

The result now follows from Lemma A. □

Next, we prove Theorem 3.3. Recall that we were able to put our post-processing into a single framework given by (3.12) as an approximation to the first-order system of ODEs (3.11). This motivates the study of the following more general initial value problem

$$\begin{aligned} \boldsymbol{u}'(x) - A(x)\boldsymbol{u}(x) &= \boldsymbol{f}(x) \quad \text{for } x \in K = (a,b), \\ \boldsymbol{u}(a) &= \boldsymbol{u}_a \end{aligned} \tag{3.22}$$

where $\boldsymbol{u} : [a,b] \to \mathbb{R}^m$, for some integer $m \geq 1$, is the unknown function, and $\boldsymbol{f} : [a,b] \to \mathbb{R}^m$ is a given function. We assume that A is a given $m \times m$ matrix such that there exists a unique solution to (3.22). Observe that such a condition is satisfied for the cases we are interested in this paper.

Let $r \geq 0$ be a polynomial degree and suppose that we approximate $\boldsymbol{u}$ by the function $\boldsymbol{u}_h \in [\mathcal{P}^r(K)]^m$ defined by requiring that the equation

$$-(\boldsymbol{u}_h, \boldsymbol{v}')_K + \boldsymbol{u}_h(b^-) \cdot \boldsymbol{v}(b^-) - (A\boldsymbol{u}_h, \boldsymbol{v})_K = (\boldsymbol{f}^*, \boldsymbol{v})_K + \boldsymbol{u}_a^* \cdot \boldsymbol{v}(a^+) \tag{3.23}$$

holds for all $\boldsymbol{v} \in [\mathcal{P}^r(K)]^m$. Here, $\boldsymbol{f}^*$ is an approximation to $\boldsymbol{f}$ such that

$$\|\boldsymbol{f} - \boldsymbol{f}^*\|_{0,K} \leq C\, h_K^{r+1}, \tag{3.24}$$

and $\boldsymbol{u}_a^*$ is an approximation to $\boldsymbol{u}_a$ such that

$$|\boldsymbol{u}_a - \boldsymbol{u}_a^*| \leq C\, h_K^{r+1} \tag{3.25}$$

where $h_K = b - a$. The magnitude of the vector $\boldsymbol{v} \in \mathbb{R}^m$ is denoted by $|\boldsymbol{v}|$, and we have extended the definitions of Sobolev norms and seminorms to vector-valued functions in an obvious fashion. We assume that the matrix A is such that the method (3.23) defines a unique solution. We also suppose that all the components of the matrix A, and of the vector-valued functions $\boldsymbol{f}$ and $\boldsymbol{f}^*$ are in $H^{r+1}(K)$.

It is not difficult to see that the proof of Theorem 3.3 follows from a successive application of the following theorem which provides an optimal error estimate for the method defined by (3.23).

Theorem 3.5. *Suppose that we approximate the solution of the initial value problem* (3.22) *by the method* (3.23)*. Then, for sufficiently small h_K, we have the error estimate*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,K} \leq C h_K^{r+1} \tag{3.26}$$

where C is a constant independent of h_K.

Remark. More general DG methods were introduced and analyzed for the initial value problem (3.22) by Delfour *et al.* in [57]. They have proved optimal error estimates as in (3.26). The same problem has also been studied by Erikkson *et al.* in [53], and by Thomée in [54]. They have proved optimal L^∞ error estimates under more restrictive regularity requirements. Moreover, their analysis is restricted to symmetric and positive definite A.

Remark. Observe that the method (3.23) differs from those studied in [57, 53, 54] in the sense that we have to use *approximate data* $\boldsymbol{f}^*$ and $\boldsymbol{u}_a^*$ since this is precisely what we need for our purposes. Moreover, the analysis we provide in this paper is significantly different from the ones that have appeared in the literature. More explicitly, we employ projection operators tailored to the special structure of the method.

Next we describe these projection operators. For any $\psi \in H^1(K)$, the function $\pi^\pm \psi \in \mathcal{P}^r(K)$ is defined on the interval $K = [a, b]$ by

$$(\psi - \pi^\pm \psi, v)_K = 0 \qquad \forall v \in \mathcal{P}^{r-1}(K), \quad \text{if } r > 0, \tag{3.27a}$$

$$(\pi^- \psi)(b^-) = \psi(b^-), \qquad (\pi^+ \psi)(a^+) = \psi(a^+). \tag{3.27b}$$

The projection operators $\boldsymbol{\Pi}^{\pm}$ acting on vector-valued functions $\boldsymbol{\psi} : K \to \mathbb{R}^m$ are defined by (3.27) applied to each component function. Notwithstanding the fact that these projection operators have been widely used for the analysis of DG methods applied to various problems, [14, 16, 51, 58, 59, 60, 28, 61, 63] in our analysis we uncover a new superconvergence property of the projection of the error which, to the best of our knowledge, has not appeared in the literature for the analysis of DG methods for the initial value problem (3.22).

The approximation properties of $\boldsymbol{\Pi}^{\pm}$, namely, that there exists a constant C independent of $\boldsymbol{\psi}$ such that

$$\left\|\boldsymbol{\psi} - \boldsymbol{\Pi}^{\pm}\boldsymbol{\psi}\right\|_{0,K} \leq C h_k^{s+1} |\boldsymbol{\psi}|_{s+1,K} \tag{3.28}$$

for any $s \in [0, r]$, can be found in the references cited above. Theorem 3.5 follows from the above approximation property, the triangle inequality

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{0,K} \leq \left\|\boldsymbol{u} - \boldsymbol{\Pi}^{-}\boldsymbol{u}\right\|_{0,K} + \left\|\boldsymbol{\Pi}^{-}\boldsymbol{u} - \boldsymbol{u}_h\right\|_{0,K},$$

and the following superconvergence result for $\boldsymbol{\Pi}^{-}\boldsymbol{e}_{\boldsymbol{u}}$.

Theorem 3.6. *Suppose that h_K is sufficiently small. Then, we have that*

$$\left\|\boldsymbol{\Pi}^{-}\boldsymbol{e}_{\boldsymbol{u}}\right\|_{0,K} \leq C\, h_K^{r+3/2} \tag{3.29}$$

where C is a constant which is independent of h_K. Moreover, if

$$|\boldsymbol{u}_a - \boldsymbol{u}_a^*| \leq C h_K^{r+3/2}, \qquad \textit{or} \qquad \boldsymbol{u}_a = \boldsymbol{u}_a^*, \tag{3.30}$$

then

$$\left\|\boldsymbol{\Pi}^{-}\boldsymbol{e}_{\boldsymbol{u}}\right\|_{0,K} \leq C\, h_K^{r+2}. \tag{3.31}$$

The proof of this theorem will be based on a duality argument. We thus begin with introducing the dual problem for any given $\boldsymbol{\eta} : K = [a, b] \to \mathbb{R}^m$ in $L^2(K)$:

$$\boldsymbol{\psi}' + A^T \boldsymbol{\psi} = \boldsymbol{\eta} \quad \text{in } K, \tag{3.32a}$$

$$\boldsymbol{\psi}(b) = \mathbf{0}. \tag{3.32b}$$

We have the following regularity for the solution of this problem.

Lemma 3.7. *Let $\boldsymbol{\psi}$ be the solution of* (4.10)*. Then*

$$|\boldsymbol{\psi}|_{1,K} + \frac{1}{h_K} \|\boldsymbol{\psi}\|_{0,K} \;\le\; C \|\boldsymbol{\eta}\|_{0,K}\,, \tag{3.33}$$

where the constant C is independent of the datum $\boldsymbol{\eta}$.

Proof. By the basic theory of first order linear systems of differential equations we have, for any $\sigma \in [a, b]$, that

$$\boldsymbol{\psi}(x) = \boldsymbol{\Psi}(x)\boldsymbol{\Psi}^{-1}(\sigma)\,\boldsymbol{\psi}(\sigma) + \boldsymbol{\Psi}(x) \int_\sigma^x \boldsymbol{\Psi}^{-1}(s)\boldsymbol{\eta}(s)\,ds$$

where $\boldsymbol{\Psi}(\cdot)$ is the fundamental matrix associated with $-A^T$. Thus, due to the zero boundary condition at $x = b$, (4.10b),

$$\boldsymbol{\psi}(x) = \boldsymbol{\Psi}(x) \int_b^x \boldsymbol{\Psi}^{-1}(s)\boldsymbol{\eta}(s)\,ds.$$

The boundedness of $\boldsymbol{\Psi}$ and $\boldsymbol{\Psi}'$ imply

$$|\boldsymbol{\psi}|_{1,K} \le C|\boldsymbol{G}|_{1,K} \quad \text{and} \quad \|\boldsymbol{\psi}\|_{0,K} \le C \|\boldsymbol{G}\|_{0,K}$$

where $\boldsymbol{G} := \int_b^x \boldsymbol{\eta}(s)\,ds$. The first part of the regularity estimate (4.11) then follows from the fact that $|\boldsymbol{G}|_{1,K} = \|\boldsymbol{G}'\|_{0,K} = \|\boldsymbol{\eta}\|_{0,K}$. To prove the second part, we get, by a simple

application of Cauchy-Schwarz inequality that

$$\begin{aligned}\|\boldsymbol{\psi}\|_{0,K}^2 \leq C\,\|\boldsymbol{G}\|_{0,K}^2 &= C\int_a^b \left[\int_b^x \boldsymbol{\eta}(s)ds\right]^2 dx \\ &\leq C\int_a^b \Big|\int_b^x ds\Big|\Big|\int_b^x |\boldsymbol{\eta}(s)|^2 ds\Big|\,dx \\ &\leq C\,h_K\,\|\boldsymbol{\eta}\|_{0,K}^2 \int_a^b dx \\ &= C\,h_K^2\,\|\boldsymbol{\eta}\|_{0,K}^2\,.\end{aligned}$$

Hence, $\|\boldsymbol{\psi}\|_{0,K} \leq C\,h_K\,\|\boldsymbol{\eta}\|_{0,K}$. This finishes the proof. □

As expected, one of the main ingredients of our error analysis is an error equation. Inserting the exact solution $\boldsymbol{u}$ of (3.22) into the DG formulation (3.23) we get

$$-(e_{\boldsymbol{u}}, \boldsymbol{v}')_K + e_{\boldsymbol{u}}(b^-)\cdot\boldsymbol{v}(b^-) - (Ae_{\boldsymbol{u}}, \boldsymbol{v})_K = (\boldsymbol{f}-\boldsymbol{f}^*, \boldsymbol{v})_K + (\boldsymbol{u}_a - \boldsymbol{u}_a^*)\cdot\boldsymbol{v}(a^+) \tag{3.34}$$

for all $\boldsymbol{v} \in [\mathcal{P}^r(K)]^m$. Note that the quantity on the right-hand side can be viewed as a *consistency error* due to the fact that we are approximating the solution $\boldsymbol{u}$ of (3.22) by using *approximate* data $\boldsymbol{f}^*$ and $\boldsymbol{u}_a^*$. If the data are exact, namely, $\boldsymbol{f} = \boldsymbol{f}^*$ and $\boldsymbol{u}_a^* = \boldsymbol{u}_a$ then we recover a classical Galerkin orthogonality property.

The orthogonality property (3.27a) of the projection operator $\boldsymbol{\Pi}^-$, and some simple algebraic manipulations yield an alternative form of (3.34) which is more amenable to our analysis

$$\begin{aligned}-(\boldsymbol{\Pi}^- e_{\boldsymbol{u}}, \boldsymbol{v}')_K + (\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(b^-)\cdot\boldsymbol{v}(b^-) - (A\,\boldsymbol{\xi}_{\boldsymbol{u}}^-, \boldsymbol{v})_K - (A\,\boldsymbol{\Pi}^- e_{\boldsymbol{u}}, \boldsymbol{v})_K \\ = (\boldsymbol{f}-\boldsymbol{f}^*, \boldsymbol{v})_K + (\boldsymbol{u}_a - \boldsymbol{u}_a^*)\cdot\boldsymbol{v}(a^+)\end{aligned} \tag{3.35}$$

where we have introduced the notation

$$\boldsymbol{\xi}_{\boldsymbol{u}}^{\pm} := \boldsymbol{u} - \boldsymbol{\Pi}^{\pm}\boldsymbol{u}. \tag{3.36}$$

Next, we state a technical lemma.

Lemma 3.8. *Consider the dual problem* (4.10) *and the method* (3.23) *approximating the solution of* (3.22)*. Then we have the following representation formula*

$$\begin{aligned}(\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\eta})_K = &-(A\,\boldsymbol{\xi}^-_{\boldsymbol{u}},\, \boldsymbol{\psi})_K \qquad\quad + (A\,\boldsymbol{\xi}^-_{\boldsymbol{u}},\, \boldsymbol{\xi}^+_{\boldsymbol{\psi}})_K + (A\,\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\xi}^+_{\boldsymbol{\psi}})_K \\ &-(\boldsymbol{f}-\boldsymbol{f}^*,\, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K \quad -(\boldsymbol{u}_a-\boldsymbol{u}^*_a)\cdot(\boldsymbol{\Pi}^+\boldsymbol{\psi})(a^+).\end{aligned} \tag{3.37}$$

We delay the proof of this lemma to the end of this section.

We are now ready to prove Theorem 4.4.

Proof. (Theorem 4.4) Setting $\boldsymbol{\eta} = \boldsymbol{\Pi}^- e_{\boldsymbol{u}}$ in (3.37) gives

$$\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|^2_{0,K} = \sum_{i=1}^{5} T_i \tag{3.38}$$

where

$$\begin{aligned}T_1 &= -(A\boldsymbol{\xi}^-_{\boldsymbol{u}},\, \boldsymbol{\psi})_K,\\ T_2 &= \quad (A\,\boldsymbol{\xi}^-_{\boldsymbol{u}},\, \boldsymbol{\xi}^+_{\boldsymbol{\psi}})_K,\\ T_3 &= \quad (A\,\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\xi}^+_{\boldsymbol{\psi}})_K,\\ T_4 &= -(\boldsymbol{f}-\boldsymbol{f}^*,\, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K,\\ T_5 &= -(\boldsymbol{u}_a-\boldsymbol{u}^*_a)\cdot(\boldsymbol{\Pi}^+\boldsymbol{\psi})(a^+).\end{aligned}$$

An estimate of $\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}$ now follows by estimating T_i for $i = 1, \ldots, 5$. By Cauchy-Schwarz inequality we have

$$|T_1| \;\le\; \left\|A\boldsymbol{\xi}^-_{\boldsymbol{u}}\right\|_{0,K} \|\boldsymbol{\psi}\|_{0,K} \;\le\; C\left\|\boldsymbol{\xi}^-_{\boldsymbol{u}}\right\|_{0,K} \|\boldsymbol{\psi}\|_{0,K}$$

where we have used the regularity assumption on the matrix A, namely, that all component of A are in $H^{r+1}(K)$, and hence in $L^2(K)$. By the approximation properties, (3.28), of $\boldsymbol{\Pi}^-$,

and the regularity of the dual problem, (4.11), we have that

$$\begin{aligned} |T_1| &\leq C\, h_K^{r+1} |\boldsymbol{u}|_{r+1,K} \;\cdot\; C h_K \left\|\boldsymbol{\eta}\right\|_{0,K} \\ &\leq C\, h_K^{r+2} \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K} \end{aligned} \tag{3.39}$$

where we have absorbed $|\boldsymbol{u}|_{r+1,K}$ in the constant C. Similarly,

$$\begin{aligned} |T_2| &\leq \left\|A\boldsymbol{\xi}_{\boldsymbol{u}}^-\right\|_{0,K} \left\|\boldsymbol{\xi}_{\boldsymbol{\psi}}^+\right\|_{0,K} \\ &\leq C\, h_K^{r+1} |\boldsymbol{u}|_{r+1,K} \;\cdot\; C\, h_K |\boldsymbol{\psi}|_{1,K} \\ &\leq C\, h_K^{r+2} |\boldsymbol{u}|_{r+1,K} |\boldsymbol{\psi}|_{1,K} \\ &\leq C\, h_K^{r+2} \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}, \end{aligned} \tag{3.40}$$

and

$$\begin{aligned} |T_3| &\leq \left\|A\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K} \left\|\boldsymbol{\xi}_{\boldsymbol{\psi}}^+\right\|_{0,K} \\ &\leq C \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K} \;\cdot\; C\, h_K |\boldsymbol{\psi}|_{1,K} \\ &\leq C\, h_K \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K} |\boldsymbol{\psi}|_{1,K} \\ &\leq C\, h_K \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}^2. \end{aligned} \tag{3.41}$$

Note that by the continuity of the projection operator $\boldsymbol{\Pi}^+$ and the regularity, (4.11), of the dual problem we have

$$\left\|\boldsymbol{\Pi}^+ \boldsymbol{\psi}\right\|_{0,K} \;\leq\; C \left\|\boldsymbol{\psi}\right\|_{0,K} \;\leq\; C h_K \left\|\boldsymbol{\eta}\right\|_{0,K} \;=\; C h_K \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}. \tag{3.42}$$

An estimate on T_4 now follows simply by the assumption (3.24). Indeed,

$$\begin{aligned} |T_4| &\leq \left\|\boldsymbol{f} - \boldsymbol{f}^*\right\|_{0,K} \left\|\boldsymbol{\Pi}^+ \boldsymbol{\psi}\right\|_{0,K} \\ &\leq C\, h_K^{r+1} \;\cdot\; C h_K \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K} \\ &\leq C\, h_K^{r+2} \left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}. \end{aligned} \tag{3.43}$$

To estimate T_5 we will use the inverse estimate

$$|(\boldsymbol{\Pi}^+\boldsymbol{\psi})(a^+)| \;\leq\; \left\|\boldsymbol{\Pi}^+\boldsymbol{\psi}\right\|_{L^\infty(K)} \;\leq\; Ch_K^{-1/2}\left\|\boldsymbol{\Pi}^+\boldsymbol{\psi}\right\|_{0,K}$$

which can be found, for example, in (p. 149 of) [34]. Now, using (3.42), we get

$$|(\boldsymbol{\Pi}^+\boldsymbol{\psi})(a^+)| \leq Ch_K^{1/2}\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}. \tag{3.44}$$

The estimate

$$|T_5| \leq C\, h_K^{r+3/2}\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}. \tag{3.45}$$

then follows from (3.44) and the assumption (3.25).

Inserting the estimates (3.39)–(3.41), (3.43), and (3.45) into (3.38) we obtain

$$\begin{aligned}\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}^2 &\leq Ch_K^{r+2}\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K} + Ch_K\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}^2 + Ch_K^{r+3/2}\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}\\ &\leq Ch_K^{r+3/2}\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K} + Ch_K\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}^2.\end{aligned}$$

If we assume that h_K is small enough so that $Ch_K < 1$ then

$$\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}^2 \leq Ch_K^{r+3/2}\left\|\boldsymbol{\Pi}^- e_{\boldsymbol{u}}\right\|_{0,K}$$

and the estimate (3.29) follows.

Observe that the loss of half a power of h_K is caused only by the estimate of the term T_5. In particular, if (3.30) is satisfied then we recover the one-full-order-superconvergent estimate (3.31). This finishes the proof. □

It remains to prove Lemma 4.9.

Proof. (Lemma 4.9) By the definition, (4.10a), of $\boldsymbol{\psi}$

$$\begin{aligned}(\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\eta})_K &= (\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\psi}')_K + (\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, A^T\boldsymbol{\psi})_K\\ &= (\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\psi}')_K + (A\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\psi})_K\end{aligned} \tag{3.46}$$

Let us work on the first term on the right-hand side. By (3.36) we have

$$(\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\psi}')_K = (\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, (\boldsymbol{\xi}^+_{\boldsymbol{\psi}})')_K + (\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, (\boldsymbol{\Pi}^+ \boldsymbol{\psi})')_K.$$

Integrating by parts on the first term on the right-hand side and using the definition, (3.27), of $\boldsymbol{\Pi}^+$ we get

$$\begin{aligned}(\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\psi}')_K = \quad & (\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(b^-) \cdot \boldsymbol{\xi}^+_{\boldsymbol{\psi}}(b^-) \quad - (\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(a^+) \cdot \boldsymbol{\xi}^+_{\boldsymbol{\psi}}(a^+) \\ & - ((\boldsymbol{\Pi}^- e_{\boldsymbol{u}})',\, \boldsymbol{\xi}^+_{\boldsymbol{\psi}})_K \qquad + (\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, (\boldsymbol{\Pi}^+ \boldsymbol{\psi})')_K \\ = \quad & (\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(b^-) \cdot \boldsymbol{\xi}^+_{\boldsymbol{\psi}}(b^-) \quad + (\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, (\boldsymbol{\Pi}^+ \boldsymbol{\psi})')_K.\end{aligned} \tag{3.47}$$

Taking $\boldsymbol{v} = \boldsymbol{\Pi}^+ \boldsymbol{\psi}$ in (3.35) we get

$$\begin{aligned}(\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, (\boldsymbol{\Pi}^+ \boldsymbol{\psi})')_K = \quad & (\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(b^-) \cdot (\boldsymbol{\Pi}^+ \boldsymbol{\psi})(b^-) \\ & - (A \boldsymbol{\xi}^-_{\boldsymbol{u}},\, \boldsymbol{\Pi}^+ \boldsymbol{\psi})_K \qquad - (A \boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\Pi}^+ \boldsymbol{\psi})_K \\ & - (\boldsymbol{f} - \boldsymbol{f}^*,\, \boldsymbol{\Pi}^+ \boldsymbol{\psi})_K \; - (\boldsymbol{u}_a - \boldsymbol{u}^*_a) \cdot (\boldsymbol{\Pi}^+ \boldsymbol{\psi})(a^+).\end{aligned}$$

Inserting this into (3.47) we get

$$\begin{aligned}(\boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\psi}')_K = & - (A \boldsymbol{\xi}^-_{\boldsymbol{u}},\, \boldsymbol{\Pi}^+ \boldsymbol{\psi})_K \qquad - (A \boldsymbol{\Pi}^- e_{\boldsymbol{u}},\, \boldsymbol{\Pi}^+ \boldsymbol{\psi})_K \\ & - (\boldsymbol{f} - \boldsymbol{f}^*,\, \boldsymbol{\Pi}^+ \boldsymbol{\psi})_K - (\boldsymbol{u}_a - \boldsymbol{u}^*_a) \cdot (\boldsymbol{\Pi}^+ \boldsymbol{\psi})(a^+)\end{aligned}$$

where we have used the fact that

$$\begin{aligned}(\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(b^-) \cdot \boldsymbol{\xi}^+_{\boldsymbol{\psi}}(b^-) + (\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(b^-) \cdot (\boldsymbol{\Pi}^+ \boldsymbol{\psi})(b^-) & \\ = (\boldsymbol{\Pi}^- e_{\boldsymbol{u}})(b^-) \cdot \boldsymbol{\psi}(b^-) & \qquad \text{by (3.36)} \\ = 0 & \qquad \text{by (4.10b).}\end{aligned}$$

Inserting the last identity into (3.46) we obtain

$$\begin{aligned}(\boldsymbol{\Pi}^- e_{\boldsymbol{u}}, \boldsymbol{\eta})_K =& -(A\boldsymbol{\xi}_{\boldsymbol{u}}^-, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K \quad -(A\boldsymbol{\Pi}^- e_{\boldsymbol{u}}, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K + (A\boldsymbol{\Pi}^- e_{\boldsymbol{u}}, \boldsymbol{\psi})_K \\ & -(\boldsymbol{f}-\boldsymbol{f}^*, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K - (\boldsymbol{u}_a - \boldsymbol{u}_a^*)\cdot(\boldsymbol{\Pi}^+\boldsymbol{\psi})(a^+) \\ =& -(A\boldsymbol{\xi}_{\boldsymbol{u}}^-, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K \quad +(A\boldsymbol{\Pi}^- e_{\boldsymbol{u}}, \boldsymbol{\xi}_{\boldsymbol{\psi}}^+)_K \\ & -(\boldsymbol{f}-\boldsymbol{f}^*, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K - (\boldsymbol{u}_a - \boldsymbol{u}_a^*)\cdot(\boldsymbol{\Pi}^+\boldsymbol{\psi})(a^+).\end{aligned}$$

The identity (3.37) now follows since

$$(A\boldsymbol{\xi}_{\boldsymbol{u}}^-, \boldsymbol{\Pi}^+\boldsymbol{\psi})_K = (A\boldsymbol{\xi}_{\boldsymbol{u}}^-, \boldsymbol{\psi})_K - (A\boldsymbol{\xi}_{\boldsymbol{u}}^-, \boldsymbol{\xi}_{\boldsymbol{\psi}}^+)_K$$

by (3.36). □

3.4 Numerical Results

In this section, we display numerical results verifying our theoretical finding. We verify numerically that the post-processing technique introduced in Section 3.2 results in a better approximation which converges to the exact solution with order $2k+1$ in the L^2-norm inside the elements, rather than merely at the nodes of the mesh. Finally, we show that this post-processing does not deteriorate even when the parameter d is extremely small. The fact that the original DG approximation converges with the optimal order $k+1$ in the L^2-norm and with order $2k+1$ at the nodes of the mesh have been proved and numerically verified in [42]. Thus we display only the history of convergence of the post-processed approximation.

In our experiments we consider two problems. In either problem we approximate the solution of (1.3)-(1.4) subject to homogeneous boundary conditions, namely, we take

$$w_0 = w_1 = u_0 = u_1 = \theta_0 = \theta_1 = 0.$$

In both examples we take $\kappa \equiv 1$ which corresponds to a circular arch. Although the theory has been carried out for arches with arbitrary geometry and κ can be any $L^\infty(\Omega_h)$ function which satisfies the mild restriction (2.10), we have to consider a circular arch since we need to compute the exact solution to the problem so that we can carry out a history of convergence study. We first employ the DG method defined by (2.2) with the numerical traces given by (4.2)-(2.5) which are obtained by setting

$$C_{16} = C_{25} = C_{34} = C_{43} = C_{52} = C_{61} = -1$$

for all x in $\mathcal{E}_h$, except $C_{16} = C_{25} = C_{34} = 0$ on $\partial\Omega$, and setting all the other coefficients to zero. Observe that these coefficients satisfy the conditions provided by (3.1)-(3.4), and hence the numerical traces of the DG solution are superconvergent of order $2k + 1$ by Theorem 3.1. The post-processing is then computed in an element-by-element fashion as described in Steps 1–4 of Section 3.2. The only difference between the two problems arise from the loading of the arch. In the first example we take

$$p \equiv q \equiv 1 \qquad \text{in } \Omega$$

which corresponds to an arch which is loaded uniformly in both the transverse and tangential directions. In the second example, we take

$$p \equiv 0, \qquad q \equiv d^{-2} \qquad \text{in } \Omega$$

which corresponds to a so-called *membrane arch*. It has no tangential loads and is loaded very strongly in the transverse direction. The transverse load is taken inversely proportional to the square of the thickness of the arch due to the fact that the membrane arch is well-known to become extremely *stiff* as d converges to zero, and it becomes impossible to observe

meaningful displacements unless such large transverse loads are applied. We have observed this phenomenon in our numerical experiments as well.

We display our numerical results in Tables 6 and 7. Therein k indicates the polynomial degree we used to define the DG method, and "mesh = i" means we employed a uniform mesh with 2^i elements. This also means that the post-processed approximation is a piecewise polynomial of degree at most $2k$ on each element. We display the numerical orders of convergence which are computed as follows. Let $\|\boldsymbol{e}^*(i)\|_0$ denote the $L^2(\Omega_h)$-norm of the error where a uniform mesh with 2^i elements has been employed to obtain the DG approximation and its post-processing. For brevity, rather than displaying the error for each individual unknown, we display the total error defined as

$$\|\boldsymbol{e}^*\|_0 := \left(\|e_w^*\|_0^2 + \|e_u^*\|_0^2 + \|e_\theta^*\|_0^2 + \|e_M^*\|_0^2 + \|e_N^*\|_0^2 + \|e_T^*\|_0^2\right)^{1/2}.$$

The order of convergence, r_i, at the level i is then defined as

$$r_i = \frac{\log\left(\frac{\|\boldsymbol{e}^*(i-1)\|_0}{\|\boldsymbol{e}^*(i)\|_0}\right)}{\log 2}.$$

In light of Theorem 3.3, we expect this quantity to approach $2k+1$ in the asymptotic regime. Furthermore, in order to verify that the quality of the post-processed approximation does not deteriorate as d becomes very small, we take $d = 10^{-1}$ and then decrease it down to $d = 10^{-8}$.

In Tables 6 and 7 we display our numerical results for the first and the second examples, respectively. In both cases we clearly see that the post-processed approximation converges with order $2k+1$ to the exact solution as predicted by Theorem 3.3. Moreover, these results do not deteriorate as the parameter d becomes extremely small and the convergence of the post-processed solution is robust with respect to d. This verifies the theoretically expected

Table 4: History of convergence of the post-processed DG approximation for the first problem.

		$d = 10^{-1}$		$d = 10^{-4}$		$d = 10^{-8}$	
k	mesh	$\|\boldsymbol{e}^*\|_0$	order	$\|\boldsymbol{e}^*\|_0$	order	$\|\boldsymbol{e}^*\|_0$	order
	5	3.01E-05	3.04	3.27E-06	3.06	3.27E-06	3.06
1	6	3.71E-06	3.02	4.00E-07	3.03	4.00E-07	3.03
	7	4.60E-07	3.01	4.95E-08	3.02	4.95E-08	3.02
	8	5.73E-08	3.01	6.15E-09	3.01	6.15E-09	3.01
	5	1.72E-10	4.92	1.96E-10	4.92	1.96E-10	4.92
2	6	5.57E-12	4.95	6.30E-12	4.96	6.30E-12	4.96
	7	1.78E-13	4.97	2.00E-13	4.98	2.00E-13	4.98
	8	5.62E-15	4.98	6.28E-15	4.99	6.28E-15	4.99
	4	8.88E-14	7.19	2.43E-15	7.86	2.43E-15	7.86
3	5	6.41E-16	7.12	1.08E-17	7.81	1.08E-17	7.81
	6	4.79E-18	7.06	5.20E-20	7.70	5.20E-20	7.70
	7	3.66E-20	7.03	2.84E-22	7.52	2.84E-22	7.52

Table 5: History of convergence of the post-processed DG approximation for the second problem.

		$d = 10^{-1}$		$d = 10^{-4}$		$d = 10^{-8}$	
k	mesh	$\|\boldsymbol{e}^*\|_0$	order	$\|\boldsymbol{e}^*\|_0$	order	$\|\boldsymbol{e}^*\|_0$	order
	5	3.00E-03	3.04	2.14E-01	3.04	2.14E-01	3.04
1	6	3.69E-04	3.02	2.64E-02	3.02	2.64E-02	3.02
	7	4.58E-05	3.01	3.28E-03	3.01	3.28E-03	3.01
	8	5.70E-06	3.01	4.08E-04	3.01	4.08E-04	3.01
	5	1.14E-09	5.50	1.12E-07	4.51	1.12E-07	4.51
2	6	8.34E-11	3.78	6.84E-09	4.04	6.84E-09	4.04
	7	3.39E-12	4.62	2.68E-10	4.67	2.68E-10	4.67
	8	1.18E-13	4.84	9.25E-12	4.86	9.25E-12	4.86
	5	6.41E-14	7.11	4.99E-12	7.11	4.99E-12	7.11
3	6	4.79E-16	7.06	3.74E-14	7.06	3.74E-14	7.06
	7	3.66E-18	7.03	2.86E-16	7.03	2.86E-16	7.03
	8	2.83E-20	7.02	2.21E-18	7.02	2.21E-18	7.02

fact that the DG methods as well as their post-processing is free from shear and membrane locking. This is remarkable especially for the membrane arch since the behavior of its solution is extremely sensitive to the value of the thickness of the arch, especially for small values of the parameter d.

3.5 Conclusion

We introduced and numerically tested a remarkably efficient and inexpensive post-processing method for the DG solutions for the Naghdi arch problem. Although the DG approximation converges with order $k+1$ when polynomials of degree k are used, the post-processed approximation superconverges with order $2k+1$. The post-processing exploits the fact that the numerical traces of the DG method converge with order $2k+1$. This result holds independently of the thickness parameter d, which shows that the post-processing as well as the DG methods are free from shear and membrane locking.

4 Hybridizable DG Methods for Naghdi Arches

4.1 Introduction

In this section, we introduce a class of hybridizable discontinuous Galerkin (HDG) methods for Naghdi arches.

Classical Galerkin methods have been analyzed for circular arches (κ identically equal to a constant) in [37, 38, 39], It has been shown that the well-known remedy of using reduced integration results in locking free continuous Galerkin methods. Recently, a family of DG methods for Naghdi arches have been developed and analyzed in [42]. It has been shown that a wide class of these methods converge with optimal order and that they are free from locking. However, they suffer from the usual criticism that DG methods have *too many* degrees of freedom compared to their conforming counterparts. Secondly, although it sheds light into many aspects of the problem, the framework provided therein does not lend itself very conveniently to developing numerical methods for solving shell problems. Through our study of HDG methods for arches in this paper, we are addressing both issues. First and foremost, it is well known that [43] HDG methods are efficiently implementable since the internal degrees of freedom are eliminated and the only globally coupled unknowns are those corresponding to element faces. Secondly, the framework we provide in the present work is much simpler in the sense that the global linear system is obtained only through the enforcement of the so-called *transmission* conditions.

On the other hand, the HDG methods were introduced in [43] in the framework of second-order elliptic problems. The main feature of these methods is that their approximate solutions can be expressed in an element-by-element fashion in terms of an approximate trace satisfying

a global weak formulation. Since the associated matrix is symmetric and positive definite, these methods can be efficiently implemented. In [44], the single-face HDG method (SFH) for second order elliptic problems was introduced. It was proved that by using polynomials of degree $k \geq 0$ for both the potential as well as the flux, the order of convergence in L^2 of both unknowns is $k+1$. Later it was shown [45] that many other DG methods, including a wide class of HDG methods, have these optimal convergence properties as well.

The methods that we develop in the present paper are extensions of those developed and analyzed in [19, 46]. This is a necessary intermediate step towards the challenging goal of designing efficient HDG methods for shells. Since the Naghdi arch model that we study here can be obtained from the two dimensional Naghdi shell model by dimensional reduction, that the structure of the methods that we describe here provides us with a framework that we can use for developing HDG methods for shells.

Finally, let us note that the extension of HDG methods for Timoshenko beams to those for the Naghdi arch model is not merely a matter of dealing with more variables. The fundamental difference herein is the fact that some of the unknowns involved, namely, the membrane stress N and the shear force T are coupled as well as the transverse and tangential displacements w and u. This coupling introduces additional technical difficulties to the analysis of the methods. The manner in which we overcome these difficulties provides us with a list of recipes to overcome those that we will most likely encounter when stepping up from HDG methods for plates to HDG methods for shells. Although the transition from beams to arches was not a trivial task, it is encouraging to see that the structure laid out in [19, 46] lends itself to a generalization to this problem.

4.2 The HDG methods

Let us describe the HDG methods under consideration. We begin by introducing our notation. To each partition of the domain Ω, we set

$$\Omega_h := \{I_j = (x_{j-1}, x_j) : \ 0 = x_0 < x_1 < \cdots < x_{R-1} < x_R = 1\}.$$

We associate the set of nodes, $\mathcal{E}_h := \{x_0, x_1, \ldots, x_R\}$, and the set of interior nodes $\mathcal{E}_h^\circ := \mathcal{E}_h \backslash \partial\Omega$; we also set $\partial\Omega_h := \{\partial K : K \in \Omega_h\}$. For each element $K \in \Omega_h$, let h_K denote the length of K, and set $h := \max_{K\in\Omega_h}\{h_K\}$. Finally, for any given polynomial degree $k \geq 0$ and an element $K \in \Omega_h$, we define $\mathcal{P}^k(K)$ as the set of polynomials of degree less than or equal to k on K. The space of piecewise polynomials of degree k on Ω is defined accordingly as

$$V_h^k := \{v : \Omega_h \mapsto \mathbb{R} : \ v|_K \in \mathcal{P}^k(K) \text{ for all } K \in \Omega_h\} \quad \text{and} \quad \boldsymbol{V}_h^k := [V_h^k]^6.$$

We also set

$$L_0^2(\mathcal{E}_h) := \{m \in L^2(\mathcal{E}_h) : m = 0 \text{ on } \partial\Omega\} \quad \text{and} \quad \boldsymbol{L}_0^2(\mathcal{E}_h) := [L_0^2(\mathcal{E}_h)]^3.$$

The HDG methods seek an approximation

$$(T_h, N_h, M_h, \theta_h, u_h, w_h, \widehat{M}_h, \widehat{u}_h, \widehat{w}_h)$$

to the exact solution

$$(T, N, M, \theta, u, w, M|_{\mathcal{E}_h}, u|_{\mathcal{E}_h}, w|_{\mathcal{E}_h}),$$

in the finite dimensional space $\boldsymbol{V}_h^k \times \boldsymbol{L}^2(\mathcal{E}_h)$. It is determined by requiring that

$$
\begin{aligned}
-(w_h, v_1') &+ \langle \widehat{w}_h, v_1\, n\rangle + (\theta_h, v_1) + (\kappa\, u_h, v_1) &&= d^2(T_h, v_1), && \text{(4.1a)}\\
-(u_h, v_2') &+ \langle \widehat{u}_h, v_2\, n\rangle - (\kappa\, w_h, v_2) &&= d^2(N_h, v_2), && \text{(4.1b)}\\
-(\theta_h, v_3') &+ \langle \widehat{\theta}_h, v_3\, n\rangle &&= (M_h, v_3), && \text{(4.1c)}\\
-(M_h, v_4') &+ \langle \widehat{M}_h, v_4\, n\rangle &&= (T_h, v_4), && \text{(4.1d)}\\
-(N_h, v_5') &+ \langle \widehat{N}_h, v_5\, n\rangle - (\kappa\, T_h, v_5) &&= (p, v_5), && \text{(4.1e)}\\
-(T_h, v_6') &+ \langle \widehat{T}_h, v_6\, n\rangle + (\kappa\, N_h, v_6) &&= (q, v_6), && \text{(4.1f)}\\
&\langle \widehat{\theta}_h, \mathsf{m}\, n\rangle &&= \langle \theta_N, \mathsf{m}\, n\rangle_{\partial\Omega}, && \text{(4.1g)}\\
&\langle \widehat{N}_h, \mathsf{u}\, n\rangle &&= 0, && \text{(4.1h)}\\
&\langle \widehat{T}_h, \mathsf{w}\, n\rangle &&= 0, && \text{(4.1i)}
\end{aligned}
$$

hold for all

$$(v_1, v_2, v_3, v_4, v_5, v_6, \mathsf{m}, \mathsf{u}, \mathsf{w}) \in \boldsymbol{V}_h^k \times L^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h).$$

Here, the outward unit normal vectors are $n(x^{\mp}) := \pm 1$ for $x \in \mathcal{E}_h$. The "volume" inner product is defined as

$$(z, v) := \sum_{K \in \Omega_h} (z, v)_K \quad \text{where} \quad (z, v)_K := \int_K z(x) v(x)\, dx,$$

and the boundary inner product is defined as

$$\langle z, v\, n\rangle := \sum_{K \in \Omega_h} \langle z, v\, n\rangle_{\partial K} \quad \text{where} \quad \langle z, v\rangle_{\partial K} := z(x_j^-) v(x_j^-) + z(x_{j-1}^+) v(x_{j-1}^+),$$

when $K = (x_{j-1}, x_j)$, and $z(x^{\pm}) := \lim_{\epsilon \downarrow 0} z(x \pm \epsilon)$ for $x \in \mathcal{E}_h$.

Note that the boundary condition on θ is imposed by equation (4.1g). The boundary conditions on w and u, respectively, are imposed as follows:

$$\widehat{w}_h = w_D \quad \text{on } \partial\Omega, \tag{4.2a}$$

$$\widehat{u}_h = u_D \quad \text{on } \partial\Omega. \tag{4.2b}$$

To complete the definition of the HDG method, we need to express the numerical traces $\widehat{T}_h$, $\widehat{N}_h$, and $\widehat{\theta}_h$ in terms of the unknowns:

$$\begin{bmatrix} \widehat{\theta}_h \\ \widehat{N}_h \\ \widehat{T}_h \end{bmatrix} = \begin{bmatrix} \theta_h \\ N_h \\ T_h \end{bmatrix} - \mathtt{S} \begin{bmatrix} M_h - \widehat{M}_h \\ u_h - \widehat{u}_h \\ w_h - \widehat{w}_h \end{bmatrix} n \tag{4.2c}$$

where

$$\mathtt{S} := \begin{bmatrix} \alpha_\theta & \tau_1 & \tau_2 \\ -\tau_1 & \alpha_N & \tau_3 \\ -\tau_2 & -\tau_3 & \alpha_T \end{bmatrix}$$

is the so-called stabilization function which is defined on $\partial\Omega_h$. Its components have to be chosen suitably to guarantee the existence and uniqueness of the approximate solution. Their choice also affects the accuracy of the method.

4.3 Existence and uniqueness of the HDG solution

In this section we provide sufficient conditions under which the HDG method introduced in the previous section defines a unique solution. As is usual for DG methods, the existence and uniqueness of the approximation depends strongly on the definition of the numerical traces (4.2). We state our existence and uniqueness result in the following theorem.

Theorem 4.1. *Consider the HDG method defined by the weak formulation* (4.1)*, and the formulas* (4.2) *for the numerical traces. Let* $\overline{\kappa}_j$ *denote the average value of* κ *on* I_j*. Suppose that*

$$h_j \leq \frac{1}{2\|\kappa - \overline{\kappa}_j\|_{L^\infty(I_j)}} \tag{4.3}$$

on the elements I_j *where* κ *is not identically equal to a constant. Suppose that the stabilization functions*

$$\alpha_T,\, \alpha_N > 0, \quad \textit{and} \quad \alpha_\theta \geq 0, \quad \textit{on} \quad \partial\Omega_h. \tag{4.4}$$

Then, for $k \geq 1$*, the method has a unique solution. For* $k = 0$*, the method defines a unique solution provided (in addition to the condition* (4.4)*) that*

$$\alpha_\theta > 0 \quad \textit{on at least one point of } \partial\Omega_h. \tag{4.5}$$

We see from (4.4) that α_T and α_N play a more important role than α_θ. Although we required the strict positivity of α_T and α_N at both ends of each element of Ω_h, the positivity of α_θ only at one end point of only one element is sufficient for the existence and uniqueness. Furthermore, this condition is needed only when $k = 0$.

There are no positivity requirements on τ_1, τ_2, or τ_3. However, a (skew) symmetry condition is implicitly imbedded into the stabilization function `S`.

Although the assumption (4.3) is a restriction on the mesh Ω_h, it can be viewed as a very mild restriction on the geometry of the arch. It basically states that, within each element, the curvature of the arch is approximately equal to that of a circle. Clearly, this is a very reasonable assumption for all practical purposes.

We prove Theorem 4.1 in Sec. 4.6.

4.4 Characterization of the approximate solution

In this section, we show that the *only* globally coupled unknowns of the HDG method defined by the weak formulation (4.1), and the formulas (4.2) for the numerical traces are the approximations at the nodes to the transverse and tangential displacement, and bending moment given by the numerical traces $\widehat{w}_h$, $\widehat{u}_h$, and $\widehat{M}_h$, respectively. We also show that the remaining components of the approximate solution can be expressed solely in terms of element-by-element-defined operators acting on $\widehat{w}_h$, $\widehat{u}_h$, and $\widehat{M}_h$. To do this, we follow the framework provided in [43] and [19].

We begin by introducing the above-mentioned locally defined operators which we call the *local solvers* associated with the method.

The first local solver is defined on the element $K \in \Omega_h$ as the mapping

$$\omega \in L^2(\partial K) \mapsto (\mathcal{T}\omega, \mathcal{N}\omega, \mathcal{M}\omega, \Theta\omega, \mathcal{U}\omega, \mathcal{W}\omega) \in \boldsymbol{\mathcal{P}}^k(K)$$

where

$$\begin{aligned}
-(\mathcal{W}\omega, v_1')_K + \langle \omega, v_1 n\rangle_{\partial K} \quad &+ (\Theta\omega, v_1)_K + (\kappa\,\mathcal{U}\omega, v_1)_K &&= d^2(\mathcal{T}\omega, v_1)_K,\\
-(\mathcal{U}\omega, v_2')_K \quad &&- (\kappa\,\mathcal{W}\omega, v_2)_K &= d^2(\mathcal{N}\omega, v_2)_K,\\
-(\Theta\omega, v_3')_K + \langle \widehat{\Theta}\omega, v_3 n\rangle_{\partial K} \quad &&&= (\mathcal{M}\omega, v_3)_K,\\
-(\mathcal{M}\omega, v_4')_K \quad &&&= (\mathcal{T}\omega, v_4)_K,\\
-(\mathcal{N}\omega, v_5')_K + \langle \widehat{\mathcal{N}}\omega, v_5 n\rangle_{\partial K} \quad &&- (\kappa\,\mathcal{T}\omega, v_5)_K &= 0,\\
-(\mathcal{T}\omega, v_6')_K + \langle \widehat{\mathcal{T}}\omega, v_6 n\rangle_{\partial K} \quad &&+ (\kappa\,\mathcal{N}\omega, v_6)_K &= 0,
\end{aligned}$$

for all $v_i \in \mathcal{P}^k(K)$ for $i = 1, \ldots, 6$. Here,

$$\begin{bmatrix} \widehat{\Theta} p \\ \widehat{\mathcal{N}} p \\ \widehat{\mathcal{T}} p \end{bmatrix} = \begin{bmatrix} \Theta p \\ \mathcal{N} p \\ \mathcal{T} p \end{bmatrix} - \mathtt{S} \begin{bmatrix} \mathcal{M} p \\ \mathcal{U} p \\ \mathcal{W} p \end{bmatrix} n$$

The second local solver is defined on the element $K \in \Omega_h$ as the mapping

$$u \in L^2(\partial K) \mapsto (\mathcal{T}u, \mathcal{U}u, \mathcal{M}u, \Theta u, \mathcal{U}u, \mathcal{W}u) \in \boldsymbol{\mathcal{P}}^k(K)$$

where

$$\begin{aligned}
-(\mathcal{W}u, v_1')_K \qquad\qquad &+(\Theta u, v_1)_K + (\kappa\,\mathcal{U}u, v_1)_K &&= d^2(\mathcal{T}u, v_1)_K, \\
-(\mathcal{U}u, v_2')_K + \langle u, v_2 n\rangle_{\partial K} \qquad\qquad &-(\kappa\,\mathcal{W}u, v_2)_K &&= d^2(\mathcal{N}u, v_2)_K, \\
-(\Theta u, v_3')_K + \langle \widehat{\Theta} u, v_3 n\rangle_{\partial K} \qquad\qquad & &&= (\mathcal{M}u, v_3)_K, \\
-(\mathcal{M}u, v_4')_K \qquad\qquad & &&= (\mathcal{T}u, v_4)_K, \\
-(\mathcal{N}u, v_5')_K + \langle \widehat{\mathcal{N}} u, v_5 n\rangle_{\partial K} \qquad\qquad &-(\kappa\,\mathcal{T}u, v_5)_K &&= 0, \\
-(\mathcal{T}u, v_6')_K + \langle \widehat{\mathcal{T}} u, v_6 n\rangle_{\partial K} \qquad\qquad &+(\kappa\,\mathcal{N}u, v_6)_K &&= 0,
\end{aligned}$$

for all $v_i \in \mathcal{P}^k(K)$ for $i = 1, \ldots, 6$. Here,

$$\begin{bmatrix} \widehat{\Theta} u \\ \widehat{\mathcal{N}} u \\ \widehat{\mathcal{T}} u \end{bmatrix} = \begin{bmatrix} \Theta u \\ \mathcal{N} u \\ \mathcal{T} u \end{bmatrix} - \mathtt{S} \begin{bmatrix} \mathcal{M} u \\ \mathcal{U} u - u \\ \mathcal{W} u \end{bmatrix} n$$

The third local solver is defined on the element $K \in \Omega_h$ as the mapping

$$\mu \in L^2(\partial K) \mapsto (\mathcal{T}\mu, \mathcal{N}\mu, \mathcal{M}\mu, \Theta\mu, \mathcal{U}\mu, \mathcal{W}\mu) \in \boldsymbol{\mathcal{P}}^k(K)$$

where

$$
\begin{aligned}
-(\mathcal{W}\mu, v_1')_K &+ (\Theta\mu, v_1)_K + (\kappa\,\mathcal{U}\mu, v_1)_K &&= d^2(\mathcal{T}\mu, v_1)_K,\\
-(\mathcal{U}\mu, v_2')_K &- (\kappa\,\mathcal{W}\mu, v_2)_K &&= d^2(\mathcal{N}\mu, v_2)_K,\\
-(\Theta\mu, v_3')_K &+ \langle\widehat{\Theta}\mu, v_3 n\rangle_{\partial K} &&= (\mathcal{M}\mu, v_3)_K,\\
-(\mathcal{M}\mu, v_4')_K &+ \langle\mu, v_4 n\rangle_{\partial K} &&= (\mathcal{T}\mu, v_4)_K,\\
-(\mathcal{N}\mu, v_5')_K &+ \langle\widehat{\mathcal{N}}\mu, v_5 n\rangle_{\partial K} &&- (\kappa\,\mathcal{T}\mu, v_5)_K = 0,\\
-(\mathcal{T}\mu, v_6')_K &+ \langle\widehat{\mathcal{T}}\mu, v_6 n\rangle_{\partial K} &&+ (\kappa\,\mathcal{N}\mu, v_6)_K = 0,
\end{aligned}
$$

for all $v_i \in \mathcal{P}^k(K)$ for $i = 1, \ldots, 6$. Here,

$$
\begin{bmatrix} \widehat{\Theta}\mu \\ \widehat{\mathcal{N}}\mu \\ \widehat{\mathcal{T}}\mu \end{bmatrix} = \begin{bmatrix} \Theta\mu \\ \mathcal{N}\mu \\ \mathcal{T}\mu \end{bmatrix} - \mathtt{S} \begin{bmatrix} \mathcal{M}\mu - \mu \\ \mathcal{U}\mu \\ \mathcal{W}\mu \end{bmatrix} n
$$

The fourth local solver is defined on the element $K \in \Omega_h$ as the mapping

$$
p \in L^2(K) \mapsto (\mathcal{T}p, \mathcal{N}p, \mathcal{M}p, \Theta p, \mathcal{U}p, \mathcal{W}p) \in \boldsymbol{\mathcal{P}}^k(K)
$$

where

$$
\begin{aligned}
-(\mathcal{W}p, v_1')_K &+ (\Theta p, v_1)_K + (\kappa\,\mathcal{U}p, v_1)_K &&= d^2(\mathcal{T}p, v_1)_K,\\
-(\mathcal{U}p, v_2')_K &- (\kappa\,\mathcal{W}p, v_2)_K &&= d^2(\mathcal{N}p, v_2)_K,\\
-(\Theta p, v_3')_K &+ \langle\widehat{\Theta}p, v_3 n\rangle_{\partial K} &&= (\mathcal{M}p, v_3)_K,\\
-(\mathcal{M}p, v_4')_K &&&= (\mathcal{T}p, v_4)_K,\\
-(\mathcal{N}p, v_5')_K &+ \langle\widehat{\mathcal{N}}p, v_5 n\rangle_{\partial K} &&- (\kappa\,\mathcal{T}p, v_5)_K = (p, v_5)_K,\\
-(\mathcal{T}p, v_6')_K &+ \langle\widehat{\mathcal{T}}p, v_6 n\rangle_{\partial K} &&+ (\kappa\,\mathcal{N}p, v_6)_K = 0,
\end{aligned}
$$

for all $v_i \in \mathcal{P}^k(K)$ for $i = 1, \ldots, 6$. Here,

$$\begin{bmatrix} \widehat{\Theta}p \\ \widehat{\mathcal{N}}p \\ \widehat{\mathcal{T}}p \end{bmatrix} = \begin{bmatrix} \Theta p \\ \mathcal{N}p \\ \mathcal{T}p \end{bmatrix} - \mathtt{S} \begin{bmatrix} \mathcal{M}p \\ \mathcal{U}p \\ \mathcal{W}p \end{bmatrix} n$$

Finally, the fifth local solver is defined on the element $K \in \Omega_h$ as the mapping

$$q \in L^2(K) \mapsto (\mathcal{T}q, \mathcal{N}q, \mathcal{M}q, \Theta q, \mathcal{U}q, \mathcal{W}q) \in \boldsymbol{\mathcal{P}}^k(K)$$

where

$$\begin{aligned}
-(\mathcal{W}q, v_1')_K \qquad\qquad &+(\Theta q, v_1)_K + (\kappa\,\mathcal{U}q, v_1)_K &&= d^2(\mathcal{T}q, v_1)_K, \\
-(\mathcal{U}q, v_2')_K \qquad\qquad &\qquad\qquad - (\kappa\,\mathcal{W}q, v_2)_K &&= d^2(\mathcal{N}q, v_2)_K, \\
-(\Theta q, v_3')_K + \langle\widehat{\Theta}q, v_3 n\rangle_{\partial K} &\qquad\qquad &&= (\mathcal{M}q, v_3)_K, \\
-(\mathcal{M}q, v_4')_K \qquad\qquad &\qquad\qquad &&= (\mathcal{T}q, v_4)_K, \\
-(\mathcal{N}q, v_5')_K + \langle\widehat{\mathcal{N}}q, v_5 n\rangle_{\partial K} &\qquad\qquad - (\kappa\,\mathcal{T}q, v_5)_K &&= 0, \\
-(\mathcal{T}q, v_6')_K + \langle\widehat{\mathcal{T}}q, v_6 n\rangle_{\partial K} &\qquad\qquad + (\kappa\,\mathcal{N}q, v_6)_K &&= (q, v_6)_K,
\end{aligned}$$

for all $v_i \in \mathcal{P}^k(K)$ for $i = 1, \ldots, 6$. Here,

$$\begin{bmatrix} \widehat{\Theta}q \\ \widehat{\mathcal{N}}q \\ \widehat{\mathcal{T}}q \end{bmatrix} = \begin{bmatrix} \Theta q \\ \mathcal{N}q \\ \mathcal{T}q \end{bmatrix} - \mathtt{S} \begin{bmatrix} \mathcal{M}q \\ \mathcal{U}q \\ \mathcal{W}q \end{bmatrix} n.$$

The function w_D, as well as any other function defined only on $\partial\Omega$ is extended to $\mathcal{E}_h$ by zero. We also set

$$\omega_h := \begin{cases} \widehat{w}_h & \text{on } \partial\Omega_h \backslash \partial\Omega \\ 0 & \text{on } \partial\Omega, \end{cases} \qquad u_h := \begin{cases} \widehat{u}_h & \text{on } \partial\Omega_h \backslash \partial\Omega \\ 0 & \text{on } \partial\Omega, \end{cases}$$

so that we have that $\widehat{w}_h = \omega_h + w_D$ and that $\widehat{u}_h = u_h + u_D$ where $\omega_h, u_h \in L^2_0(\mathcal{E}_h)$. Also, to simplify the notation we write $\mu_h := \widehat{M}_h$ on $\partial\Omega_h$. We can now state a characterization of the approximate solution in terms of the local solvers.

Theorem 4.2. *Suppose that the conditions of Theorem 4.1 are satisfied. Then the approximate solution* $(T_h,\, N_h,\, M_h,\, \theta_h,\, u_h,\, w_h,\, \mu_h,\, u_h,\, \omega_h) \in \boldsymbol{V}_h^k \times L^2(\mathcal{E}_h) \times L^2_0(\mathcal{E}_h) \times L^2_0(\mathcal{E}_h)$ *given by the HDG method can be expressed in terms of the local solvers as*

$$
\begin{aligned}
T_h &= \mathcal{T}\omega_h + \mathcal{T}w_D + \mathcal{T}u_h + \mathcal{T}u_D + \mathcal{T}\mu_h + \mathcal{T}p + \mathcal{T}q,\\
N_h &= \mathcal{N}\omega_h + \mathcal{N}w_D + \mathcal{N}u_h + \mathcal{N}u_D + \mathcal{N}\mu_h + \mathcal{N}p + \mathcal{N}q,\\
M_h &= \mathcal{M}\omega_h + \mathcal{M}w_D + \mathcal{M}u_h + \mathcal{M}u_D + \mathcal{M}\mu_h + \mathcal{M}p + \mathcal{M}q,\\
\theta_h &= \Theta\omega_h + \Theta w_D + \Theta u_h + \Theta u_D + \Theta\mu_h + \Theta p + \Theta q,\\
u_h &= \mathcal{U}\omega_h + \mathcal{U}w_D + \mathcal{U}u_h + \mathcal{U}u_D + \mathcal{U}\mu_h + \mathcal{U}p + \mathcal{U}q,\\
w_h &= \mathcal{W}\omega_h + \mathcal{W}w_D + \mathcal{W}u_h + \mathcal{W}u_D + \mathcal{W}\mu_h + \mathcal{W}p + \mathcal{W}q,
\end{aligned}
$$

where $(\mu_h, u_h, \omega_h) \in L^2(\mathcal{E}_h) \times L^2_0(\mathcal{E}_h) \times L^2_0(\mathcal{E}_h)$ *satisfies*

$$
a_h(\mu_h, u_h, \omega_h; \mathsf{m}, \mathsf{u}, \mathsf{w}) = \ell_h(\mathsf{m}, \mathsf{u}, \mathsf{w})
$$

for all $(\mathsf{m}, \mathsf{u}, \mathsf{w}) \in L^2(\mathcal{E}_h) \times L^2_0(\mathcal{E}_h) \times L^2_0(\mathcal{E}_h)$. *Here,* a_h *and* ℓ_h *are suitably defined* bilinear *and* linear *forms, respectively.*

Explicit expressions for a_h and ℓ_h as well as the proof of the above Theorem can be found in Appendix **B**. Let us remark, however, that they are obtained by a suitable rewriting of the *conservativity* conditions (4.1g)-(4.1i), see [43, 19].

Note that the total number of globally coupled unknowns in the equation in Theorem 4.2 is $3R-1$ where R is the number of elements in Ω_h. In particular, it is independent of the

polynomial degree k. This should be contrasted with the total number of globally coupled unknowns for classical DG methods [42] for the same problem, namely, $6R(k+1)$. Thus, the HDG method has significantly less number of unknowns than its DG counterpart. This is what we mean when we say that HDG methods are *efficiently implementable*.

4.5 Main Results

In this section, we present our main results. Detailed proofs of these results will be provided in Sec. 4.6. This section is organized as follows. We begin with defining a new projection operator tailored to the structure of the numerical traces of the HDG method. Subsequently, we state a theorem displaying the approximation properties of this new projection. We then present a superconvergence estimate on the projection of the error which can be considered as *the* main result of the paper since the remaining results in this section, namely, a priori error estimate for the L^2-norm of the error and a superconvergence result at the nodes of the mesh are direct consequences of it. We end this section by stating the above-mentioned a priori estimate and the nodal superconvergence result.

4.5.1 The projection

We begin with introducing the main *tool* of our error analysis, namely, a new projection operator

$$\boldsymbol{\Pi} = (\Pi_T, \Pi_N, \Pi_M, \Pi_\theta, \Pi_u, \Pi_w) : \boldsymbol{H}^1(\Omega_h) \to \boldsymbol{V}_h^k,$$

associated with the HDG methods. Here, $\boldsymbol{H}^1(\Omega_h) := [H^1(\Omega_h)]^6$. This projection operator is a generalization of the one introduced in [47] for the error analysis of HDG methods for second order elliptic problems and the one introduced in [46] for that of the HDG methods

for fourth order problems. It is defined as follows. Given a function $\mathbf{z} = (z_1, \dots, z_6) \in \boldsymbol{H}^1(\Omega_h)$ and an arbitrary subinterval $K \in \Omega_h$, the restriction of $\boldsymbol{\Pi} : \boldsymbol{H}^1(\Omega_h) \to \boldsymbol{V}_h^k$, to K is defined as the element of $\boldsymbol{\mathcal{P}}^k(K)$ that satisfies

$$(\Pi_T z_1 - z_1, v_1)_K = (\Pi_N z_2 - z_2, v_2)_K = (\Pi_M z_3 - z_3, v_3)_K = 0, \tag{4.7a}$$

$$(\Pi_\theta z_4 - z_4, v_4)_K = (\Pi_u z_5 - z_5, v_5)_K = (\Pi_w z_6 - z_6, v_6)_K = 0, \tag{4.7b}$$

for all $(v_1, \dots, v_6) \in \boldsymbol{\mathcal{P}}^{k-1}(K)$, and

$$\begin{bmatrix} z_4 \\ z_2 \\ z_1 \end{bmatrix} = \begin{bmatrix} \Pi_\theta z_4 \\ \Pi_N z_2 \\ \Pi_T z_1 \end{bmatrix} - \mathtt{S} \begin{bmatrix} \Pi_\theta z_3 - z_3 \\ \Pi_N z_5 - z_5 \\ \Pi_T z_6 - z_6 \end{bmatrix} n \quad \text{on } \partial K. \tag{4.7c}$$

Note that when $k = 0$, the projection is defined solely by (4.7c). Note also that the last set of equations reflects the form of the equations (4.2) defining the numerical traces $\widehat{\theta}_h$, $\widehat{N}_h$, and $\widehat{T}_h$. As we are going to see in the next subsection, this is what allows us to obtain a very simple set of equations for the projection of the errors.

Finally, let us point out that the projection is well defined under mild conditions on the stabilization function $\mathtt{S}$. To see this, note that the total number of unknowns involved in the linear system that is needed to be solved for computing $\boldsymbol{\Pi}\mathbf{z}$ is $6(k+1)$ since each component of the projection has $k+1$ degrees of freedom. On the other hand, the total number of linearly independent equations provided by the definition of the projection is also $6(k+1)$. The existence and uniqueness of the projection then follows from the approximation properties of the projection; see below.

4.5.2 The equations for the projection of the errors

As was pointed out in the Introduction, the projection should be devised in such a way that the equations of the projection of the errors be as simple as possible. Let us show that this is indeed the case.

Let us begin with introducing some notation. We set

$$\mathbf{z} = (T, N, M, \theta, u, w), \quad \mathbf{z}_h = (T_h, N_h, M_h, \theta_h, u_h, w_h),$$

and similarly for $\widehat{\mathbf{z}}_h$. The errors are defined as

$$e_z := z - z_h, \quad \widehat{z}_h := z - \widehat{z}_h,$$

for any $z \in \{T, N, M, \theta, u, w\}$ and we set $\boldsymbol{e} := \mathbf{z} - \mathbf{z}_h$ on Ω_h, and $\widehat{\boldsymbol{e}} := \mathbf{z} - \widehat{\mathbf{z}}_h$ on $\mathcal{E}_h$. We also define $\boldsymbol{v} := (v_1, v_2, v_3, v_4, v_5, v_6)$.

Since the exact solution $\mathbf{z}$ of the governing equations (1.3) satisfies the formulation of the

HDG approximation, (4.1), we immediately see that the equations for the errors are

$$\begin{aligned}
-(e_w, v_1') &+ \langle \widehat{e}_w, v_1 n\rangle + (e_\theta, v_1) + (\kappa\, e_u, v_1) &&= d^2(e_T, v_1),\\
-(e_u, v_2') &+ \langle \widehat{e}_u, v_2 n\rangle - (\kappa\, e_w, v_2) &&= d^2(e_N, v_2),\\
-(e_\theta, v_3') &+ \langle \widehat{e}_\theta, v_3 n\rangle &&= (e_M, v_3),\\
-(e_M, v_4') &+ \langle \widehat{e}_M, v_4 n\rangle &&= (e_T, v_4),\\
-(e_N, v_5') &+ \langle \widehat{e}_N, v_5 n\rangle - (\kappa\, e_T, v_5) &&= 0,\\
-(e_T, v_6') &+ \langle \widehat{e}_T, v_6 n\rangle + (\kappa\, e_N, v_6) &&= 0,\\
&\langle \widehat{e}_\theta, \mathsf{m}\, n\rangle &&= 0,\\
&\langle \widehat{e}_N, \mathsf{u}\, n\rangle &&= 0,\\
&\langle \widehat{e}_T, \mathsf{w}\, n\rangle &&= 0,
\end{aligned}$$

hold for all

$$(\boldsymbol{v}, \mathsf{m}, \mathsf{u}, \mathsf{w}) \in \boldsymbol{V}_h^k \times L^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h).$$

Hence, defining

$$\boldsymbol{\delta} := (\delta_T, \delta_N, \delta_M, \delta_\theta, \delta_u, \delta_w) \quad \text{where} \quad \delta_z := z - \Pi_z z$$

we obtain

$$
\begin{aligned}
-(\Pi_w e_w, v_1') &+ \langle \widehat{e}_w, v_1 n\rangle + (\Pi_\theta e_\theta + \delta_\theta, v_1) + (\kappa(\Pi_u e_u + \delta_u), v_1) && \text{(4.8a)}\\
&&& - d^2(\Pi_T e_T + \delta_T, v_1) = 0,\\
-(\Pi_u e_u, v_2') &+ \langle \widehat{e}_u, v_2 n\rangle && - (\kappa(\Pi_w e_w + \delta_w), v_2) && \text{(4.8b)}\\
&&& - d^2(\Pi_N e_N + \delta_N, v_2) = 0,\\
-(\Pi_\theta e_\theta, v_3') &+ \langle \widehat{e}_\theta, v_3 n\rangle && - (\Pi_M e_M + \delta_M, v_3) = 0, && \text{(4.8c)}\\
-(\Pi_M e_M, v_4') &+ \langle \widehat{e}_M, v_4 n\rangle && - (\Pi_T e_T + \delta_T, v_4) = 0, && \text{(4.8d)}\\
-(\Pi_N e_N, v_5') &+ \langle \widehat{e}_N, v_5 n\rangle && - (\kappa(\Pi_T e_T + \delta_T), v_5) = 0, && \text{(4.8e)}\\
-(\Pi_T e_T, v_6') &+ \langle \widehat{e}_T, v_6 n\rangle && + (\kappa(\Pi_N e_N + \delta_N), v_6) = 0, && \text{(4.8f)}\\
&&& \langle \widehat{e}_\theta, \mathsf{m}\, n\rangle = 0, && \text{(4.8g)}\\
&&& \langle \widehat{e}_N, \mathsf{u}\, n\rangle = 0, && \text{(4.8h)}\\
&&& \langle \widehat{e}_T, \mathsf{w}\, n\rangle = 0, && \text{(4.8i)}
\end{aligned}
$$

for all

$$(\boldsymbol{v}, \mathsf{m}, \mathsf{u}, \mathsf{w}) \in \boldsymbol{V}_h^k \times L^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h).$$

Note that we have used the orthogonality property of the projection (4.7) in each of the first terms of the first six equations.

To complete the error equations, we have to add the boundary conditions

$$\widehat{e}_w = \widehat{e}_u = 0 \quad \text{on } \partial\Omega, \tag{4.9a}$$

as well as the equations relating the errors inside the elements to the errors of the numerical traces, namely,

$$\begin{bmatrix} \widehat{e}_\theta \\ \widehat{e}_N \\ \widehat{e}_T \end{bmatrix} = \begin{bmatrix} \Pi_\theta e_\theta \\ \Pi_N e_N \\ \Pi_T e_T \end{bmatrix} - \mathtt{S} \begin{bmatrix} \Pi_M e_M - \widehat{e}_M \\ \Pi_u e_u - \widehat{e}_u \\ \Pi_w e_w - \widehat{e}_w \end{bmatrix} n \qquad \text{on } \partial\Omega_h. \tag{4.9b}$$

These equations hold as a direct consequence of the parallelism between the definition of the numerical traces of the HDG method, (4.2c), and the definition of the projection, (4.7c).

The *simplicity* of the error equations (4.8) and (4.9) for $\mathbf{\Pi} \boldsymbol{e}$ we have been referring to resides in the fact that they differ from the HDG approximation *only* by a *volume* integral of the approximation error $\boldsymbol{\delta}$.

4.5.3 Approximation properties of the projection $\mathbf{\Pi}$

In this subsection we state a theorem displaying the approximation properties of the projection $\mathbf{\Pi}$. First, we need to introduce some notation. Let $K = (x_L, x_R)$ be an element of Ω_h. For any function z on K, we define $z^- := z(x_L), z^+ := z(x_R)$. We denote the usual norm and seminorm on a Sobolev space $H^s(D)$ by $\|\cdot\|_{s,D}$ and $|\cdot|_{s,D}$, respectively. We drop the first subindex if $s = 0$, and the second one if $D = \Omega$ or $D = \Omega_h$. We also define the seminorm of a vector-valued function $\boldsymbol{\varphi} = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6)$ as

$$|\boldsymbol{\varphi}|_{s,D} := (|\phi_1|_{s,D}^2 + \cdots + |\phi_6|_{s,D}^2)^{\frac{1}{2}}.$$

Its norm is defined similarly.

Theorem 4.3. *We have for any s in $[1, k+1]$ that*

$$\|\boldsymbol{\delta}\| \leq C\, C_{\mathtt{S}}\, h^s\, |\mathbf{z}|_s$$

Here, C is a constant independent of the discretization parameters and $\mathbf{z}$, and $C_{\mathtt{S}}$ is given by

$$C_{\mathtt{S}} := \|\mathtt{P}\|_\infty + \left\|\mathtt{P}\,\mathtt{S}^+\right\|_\infty + \left\|\mathtt{P}\,\mathtt{S}^-\right\|_\infty + \left\|\mathtt{S}^+\mathtt{P}\right\|_\infty + \left\|\mathtt{S}^-\mathtt{P}\right\|_\infty + \left\|\mathtt{S}^+\mathtt{P}\,\mathtt{S}^-\right\|_\infty$$

where $\mathtt{P} := (\mathtt{S}^+ + \mathtt{S}^-)^{-1}$ and $\|\cdot\|_\infty$ denotes the subordinate matrix norm induced by the supremum norm on the Euclidean space.

A detailed proof of this result is given in Section 4.6. Let us note that we stated the above result for the exact solution $\mathbf{z}$ merely for notational convenience. In fact, the result remains valid if we replace $\mathbf{z}$ with any $(\phi_1, \ldots, \phi_6) \in \boldsymbol{H}^{s+1}(\Omega_h)$.

Note that $C_{\mathtt{S}}$ and hence the approximation properties of $\boldsymbol{\Pi}$ depend on the choice of $\mathtt{S}$ and hence that of the functions α_θ, α_N, α_T, τ_1, τ_2, and τ_3. It is easy to see that setting all of these functions to quantities of $\mathcal{O}(1)$ we get that $C_{\mathtt{S}} = \mathcal{O}(1)$ and hence the projection converges optimally. Setting one or more of the stabilization functions to quantities of $\mathcal{O}(1/h)$ may possibly degrade the order of convergence due to the terms $\mathtt{S}^+$ and $\mathtt{S}^-$. On the other hand, if we set some of these functions to $\mathcal{O}(h)$ the order of convergence may decrease again due to the presence of the inverse term $\mathtt{P}$ in $C_{\mathtt{S}}$. It is possible to further play with the choice of $\mathtt{S}$ and find combinations such that $C_{\mathtt{S}} = \mathcal{O}(1)$ but we will not pursue this here since we already have a very simple choice for which the projection converges optimally. Finally, we would like to point out that this simple choice of $\mathcal{O}(1)$ stabilization functions is typical of HDG methods [47, 48, 49]

4.5.4 Superconvergence of the projection of the errors

Here, we present an estimate of the projection of the errors. It is stated in terms of the solution of the so-called dual problem we define next. For any given

$$\boldsymbol{\eta} := (\eta_T, \eta_N, \eta_M, \eta_\theta, \eta_u, \eta_w) \in \boldsymbol{L}^2(\Omega),$$

the function

$$\boldsymbol{\psi} := (\psi_T, \psi_N, \psi_M, \psi_\theta, \psi_u, \psi_w) \in \boldsymbol{H}^1(\Omega)$$

is the solution of the associated dual-problem

$$\psi'_w - \psi_\theta + \kappa\psi_u = d^2\psi_T + \eta_T \quad \text{in } \Omega \tag{4.10a}$$

$$\psi'_u - \kappa\psi_w = d^2\psi_N + \eta_N \quad \text{in } \Omega \tag{4.10b}$$

$$\psi'_\theta = \psi_M - \eta_M \quad \text{in } \Omega \tag{4.10c}$$

$$\psi'_M = -\psi_T + \eta_\theta \quad \text{in } \Omega \tag{4.10d}$$

$$\psi'_N - \kappa\psi_T = -\eta_u \quad \text{in } \Omega \tag{4.10e}$$

$$\psi'_T + \kappa\psi_N = -\eta_w \quad \text{in } \Omega \tag{4.10f}$$

$$\psi_w = \psi_u = \psi_\theta = 0 \quad \text{on } \partial\Omega. \tag{4.10g}$$

We assume that the solution of this problem satisfies the following elliptic regularity result:

$$\|\boldsymbol{\psi}\|_1 \leq C_{reg} \|\boldsymbol{\eta}\|, \tag{4.11}$$

where the constant C_{reg} is independent of the datum $\boldsymbol{\eta}$ and the thickness d. A proof of this regularity estimate can be given using classical techniques of the theory of linear systems of differential equations. For details, we refer to the Appendix of [46] and Lemma 4.6 of [50].

We are now ready to state a theorem which can be regarded as the main result of this paper.

Theorem 4.4. *For $k \geq 1$, we have that, if h sufficiently small,*

$$\|\mathbf{\Pi e}\| \leq C\, C_{reg} h\, \|\boldsymbol{\delta}\|\,.$$

For $k = 0$, we have

$$\|\mathbf{\Pi e}\| \leq C\, C_{reg}\, \|\boldsymbol{\delta}\|\,.$$

Here C is a constant independent of the data of the problem and of the discretization parameters.

4.5.5 A priori error estimates

Next we present an estimate for the error in HDG approximation as an immediate consequence of the last result.

Theorem 4.5. *Suppose that the exact solution $\boldsymbol{\varphi}$ of (1.3) belongs to $\boldsymbol{H}^{k+1}(\Omega_h)$. Then, for $k \geq 1$ and h sufficiently small, we have*

$$\|\boldsymbol{e}\| \leq (1 + C\, C_{reg} h)\, \|\boldsymbol{\delta}\|\,.$$

For $k = 0$, we have

$$\|\boldsymbol{e}\| \leq (1 + C\, C_{reg})\, \|\boldsymbol{\delta}\|\,.$$

Here C is a constant independent of the data of the problem and of the discretization parameters.

Note that the error estimate appearing in the above theorem shows that, if the matrix-valued function $\mathtt{S}$ is chosen is such a way that $C_{\mathtt{S}}$ is uniformly bounded, the HDG method

is optimally convergent, that is, $\|\boldsymbol{e}\| = \mathcal{O}(h^{k+1})$ for smooth solutions and it is free from shear and membrane locking. The method is locking-free because the constant $C_{\mathtt{s}}$ does not depend on the parameter d and because the seminorms appearing on the right-hand side of the estimate can be bounded uniformly with respect to d by using the techniques employed in [20].

4.5.6 Superconvergence at the nodes

Our next result states that the numerical traces of the HDG solution superconverge. To state this result we need to introduce the Green's functions associated with the problem under consideration. For any superindex $\star \in \{T, N, M, \theta, u, w\}$, and any point $y \in (0,1)$, we define

$$\boldsymbol{G}^{\star}_{y} := (G^{\star}_{T,y}, G^{\star}_{N,y}, G^{\star}_{M,y}, G^{\star}_{\theta,y}, G^{\star}_{u,y}, G^{\star}_{w,y})$$

as the solution of

$$\begin{aligned}
d\, G^{\star}_{w,y}/dx \quad & -G^{\star}_{\theta,y} + \kappa G^{\star}_{u,y} && = d^2 G^{\star}_{T,y}, \\
d\, G^{\star}_{u,y}/dx \quad & && - \kappa G^{\star}_{w,y} && = d^2 G^{\star}_{N,y}, \\
d\, G^{\star}_{\theta,y}/dx \quad & && && = G^{\star}_{M,y}, \\
d\, G^{\star}_{M,y}/dx \quad & && && = -G^{\star}_{T,y}, \\
d\, G^{\star}_{N,y}/dx \quad & && - \kappa G^{\star}_{T,y} && = 0, \\
d\, G^{\star}_{T,y}/dx \quad & && + \kappa G^{\star}_{N,y} && = 0,
\end{aligned} \tag{4.12}$$

in $(0,y) \cup (y,1)$ that satisfies the boundary conditions

$$G^{\star}_{w,y} = G^{\star}_{u,y} = G^{\star}_{\theta,y} = 0 \quad \text{on } \partial\Omega, \tag{4.13}$$

and the jump conditions

$$\begin{aligned} [\![G^\star_{w,y}]\!](y) &= -\delta_{\star T}, \qquad [\![G^\star_{u,y}]\!](y) = -\delta_{\star N}, \qquad [\![G^\star_{\theta,y}]\!](y) = \delta_{\star M}, \\ [\![G^\star_{M,y}]\!](y) &= -\delta_{\star \theta}, \qquad [\![G^\star_{N,y}]\!](y) = \delta_{\star u}, \qquad [\![G^\star_{T,y}]\!](y) = \delta_{\star w}. \end{aligned} \tag{4.14}$$

Here, $\delta_{ab} = 1$ if $a = b$ and $\delta_{ab} = 0$ otherwise. The *jump* operator, $[\![\,\cdot\,]\!]$, is defined by

$$[\![\varphi]\!](x) := \varphi(x^-) - \varphi(x^+) \quad \text{ for } x \in \mathcal{E}_h.$$

We also define, for $t \in \{0, 1\}$, $\boldsymbol{G}^\star_t = \lim_{y \to t} \boldsymbol{G}^\star_y$.

When there is no confusion, we will drop the superindex and the second subindex of the Green's function and write, for instance, G_θ instead of $G^\star_{\theta,y}$. Finally, we define

$$\boldsymbol{\delta}^z_i := (\delta_{G^z_{T,x_i}}, \delta_{G^z_{N,x_i}}, \delta_{G^z_{M,x_i}}, \delta_{G^z_{\theta,x_i}}, \delta_{G^z_{u,x_i}}, \delta_{G^z_{w,x_i}})$$

where

$$\delta_{G^z_{\phi,x_i}} = G^z_{\phi,x_i} - \Pi_\phi G^z_{\phi,x_i}$$

for $z, \phi \in \{T, N, M, \theta, u, w\}$, and $x_i \in \mathcal{E}_h$.

We are now ready to present our superconvergence result of the numerical traces.

Theorem 4.6. *Under the same assumptions as in Theorem 4.5, we have*

$$|(z - \widehat{z}_h)(x_i)| \leq C_{k-1} h^k |\mathbf{z}|_{k+1} \, \|\boldsymbol{\delta}^z_i\| + C \, \|\boldsymbol{e}\| \, \|\boldsymbol{\delta}^z_i\|$$

for $z \in \{T, N, M, \theta, u, w\}$, and $x_i \in \mathcal{E}_h$. Here C_{k-1} is a constant that depends solely on the polynomial degree k.

Note that, for any given $k \geq 0$, if κ is a smooth function in Ω_h, the exact solution $\mathbf{z}$ belongs to $\boldsymbol{H}^{k+1}(\Omega_h)$; see [36]. This regularity result is also valid for the Green's functions

since in this case we take $p = q = 0$. Hence, we can assume that $\boldsymbol{G}^z_{x_i}$ belongs to $\boldsymbol{H}^{k+1}(\Omega_h)$. As a consequence, $\|\boldsymbol{\delta}^z_i\| = \mathcal{O}(h^{k+1})$ and the above result states that, if the constant C_{s} is uniformly bounded, *all* the numerical traces superconverge with order $2k+1$ at each node. A similar result was proved for the DG methods for Timoshenko beams studied in [20] and for Naghdi arches in [42] and HDG methods for Timoshenko beams in [19, 46].

An immediate application of the superconvergence result of Theorem 4.6 is an element-by-element postprocessing of the approximate solution provided by the HDG method. *All* the six components of the postprocessed solution converge to the exact solution with order $2k+1$, not only at the nodes, but also *uniformly* at the interior of Ω_h. For details, see [50] where we carried this out in the context of classical DG methods for Naghdi arches but exactly the same postprocessing technique also works for HDG method since the postprocessing described therein is independent of how the numerical traces have been computed.

4.6 Proofs

In this section, we provide detailed proofs of the theoretical results we have stated in Secs. 4.3 and 4.5. We proceed in the order in which the results have appeared in the paper. Each subsection is devoted to the proof of one specific result.

4.6.1 Existence and uniqueness result: Proof of Theorem 4.1

In this subsection we give a proof of the existence and uniqueness theorem stated in Section 4.3. Throughout this subsection we assume that the hypothesis of Theorem 4.1, namely, (4.4) (for $k \geq 1$) and (4.5) (for $k = 0$), are satisfied. The proof of Theorem 4.1 is based on the following technical lemmas.

Lemma 4.7. *Let* $(T_h, N_h, M_h, \theta_h, u_h, w_h, \widehat{M}_h, \widehat{u}_h, \widehat{w}_h)$ *be the HDG solution defined by the weak formulation* (4.1)*, and the formulas* (4.2) *for the numerical traces. Then we have the following identity*

$$\Theta_{int} + \Theta_{tr} = \Theta_{ld} + \Theta_{bc}, \tag{4.15}$$

where

$$\begin{aligned}
\Theta_{int} &= d^2(T_h, T_h) &&+ d^2(N_h, N_h) &&+ (M_h, M_h),\\
\Theta_{tr} &= \langle \alpha_T, (w_h - \widehat{w}_h)^2\rangle &&+ \langle \alpha_N, (u_h - \widehat{u}_h)^2\rangle &&+ \langle \alpha_\theta, (M_h - \widehat{M}_h)^2\rangle,\\
\Theta_{ld} &= -(q, w_h) &&- (p, u_h), &&\\
\Theta_{bc} &= \langle w_D, \widehat{T}_h n\rangle_{\partial\Omega} &&+ \langle u_D, \widehat{N}_h n\rangle_{\partial\Omega} &&+ \langle \theta_N, \widehat{M}_h n\rangle_{\partial\Omega}.
\end{aligned}$$

Proof. Taking $v_1 = T_h$, $v_2 = N_h$, and $v_3 = M_h$ in (4.1), and adding the resulting equations, we obtain

$$\begin{aligned}
\Theta_{int} = &- (w_h, T_h') + \langle \widehat{w}_h, T_h n\rangle &&+ (\theta_h, T_h) + (\kappa u_h, T_h)\\
&- (u_h, N_h') + \langle \widehat{u}_h, N_h n\rangle &&- (\kappa w_h, N_h)\\
&- (\theta_h, M_h') + \langle \widehat{\theta}_h, M_h n\rangle.
\end{aligned} \tag{4.16}$$

Integrating by parts on the term (w_h, T_h') and using (4.1f) with $v_6 = w_h$ implies

$$-(w_h, T_h') = \langle \widehat{T}_h - T_h, w_h n\rangle + (\kappa N_h, w_h) - (q, w_h). \tag{4.17a}$$

Similarly, using (4.1e) with $v_5 = u_h$ implies

$$-(u_h, N_h') = \langle \widehat{N}_h - N_h, u_h n\rangle - (\kappa T_h, u_h) - (p, u_h), \tag{4.17b}$$

whereas (4.1d) with $v_4 = \theta_h$ implies

$$-(\theta_h, M_h') = -\langle \theta_h, M_h n\rangle + \langle \widehat{M}_h, \theta_h n\rangle - (T_h, \theta_h). \tag{4.17c}$$

Using (4.17) in (4.16) and carrying out some cancelations, we get that

$$\begin{aligned}\Theta_{int} = \Theta_{ld} &+ \langle \widehat{T}_h - T_h, w_h n\rangle \ + \langle \widehat{w}_h, T_h n\rangle \\ &+ \langle \widehat{N}_h - N_h, u_h n\rangle + \langle \widehat{u}_h, N_h n\rangle \\ &+ \langle \widehat{\theta}_h - \theta_h, M_h n\rangle \ + \langle \widehat{M}_h, \theta_h n\rangle.\end{aligned} \tag{4.18}$$

Adding and subtracting the term $\langle \widehat{w}_h, \widehat{T}_h n\rangle$, we see that

$$\langle \widehat{T}_h - T_h, w_h n\rangle + \langle \widehat{w}_h, T_h n\rangle = \langle \widehat{T}_h - T_h, (w_h - \widehat{w}_h) n\rangle + \langle \widehat{w}_h, \widehat{T}_h n\rangle, \tag{4.19a}$$

and similarly that

$$\langle \widehat{N}_h - N_h, u_h n\rangle + \langle \widehat{u}_h, N_h n\rangle = \langle \widehat{N}_h - N_h, (u_h - \widehat{u}_h) n\rangle + \langle \widehat{u}_h, \widehat{N}_h n\rangle, \tag{4.19b}$$

$$\langle \widehat{\theta}_h - \theta_h, M_h n\rangle + \langle \widehat{M}_h, M_h n\rangle = \langle \widehat{\theta}_h - \theta_h, (M_h - \widehat{M}_h) n\rangle + \langle \widehat{\theta}_h, \widehat{M}_h n\rangle. \tag{4.19c}$$

Using (4.19) in (4.18), we have

$$\begin{aligned}\Theta_{int} = \Theta_{ld} &+ \langle \widehat{T}_h - T_h, (w_h - \widehat{w}_h) n\rangle + \langle \widehat{w}_h, \widehat{T}_h n\rangle \\ &+ \langle \widehat{N}_h - N_h, (u_h - \widehat{u}_h) n\rangle + \langle \widehat{u}_h, \widehat{N}_h n\rangle \\ &+ \langle \widehat{\theta}_h - \theta_h, (M_h - \widehat{M}_h) n\rangle + \langle \widehat{\theta}_h, \widehat{M}_h n\rangle.\end{aligned} \tag{4.20}$$

By the definition of the numerical traces, (4.2), we have that

$$\begin{aligned}\langle \widehat{T}_h - T_h, (w_h - \widehat{w}_h) n\rangle &+ \langle \widehat{N}_h - N_h, (u_h - \widehat{u}_h) n\rangle \\ &+ \langle \widehat{\theta}_h - \theta_h, (M_h - \widehat{M}_h) n\rangle = -\Theta_{tr}.\end{aligned}$$

Hence, (4.20) can be written as

$$\Theta_{int} + \Theta_{tr} = \Theta_{ld} + \langle \widehat{w}_h, \widehat{T}_h n\rangle + \langle \widehat{u}_h, \widehat{N}_h n\rangle + \langle \widehat{\theta}_h, \widehat{M}_h n\rangle.$$

The result follows once we note that

$$\langle \widehat{w}_h, \widehat{T}_h n\rangle + \langle \widehat{u}_h, \widehat{N}_h n\rangle + \langle \widehat{\theta}_h, \widehat{M}_h n\rangle = \Theta_{bc}$$

by (4.1g)-(4.1i), (4.2a), and (4.2b). □

Before proving Theorem 4.1 we state and prove an auxiliary lemma in which we collect some intermediate results.

Lemma 4.8. *Consider the HDG method defined by* (4.1)*, with the formulas* (4.2) *for the numerical traces. Suppose that the data of the problem is given by*

$$p = q = 0 \text{ in } \Omega, \quad w_D = u_D = \theta_N = 0 \text{ on } \partial\Omega, \tag{4.21}$$

then

$$T_h = N_h = M_h = 0 \quad \text{in } \Omega_h, \tag{4.22a}$$

$$\widehat{w}_h - w_h = \widehat{u}_h - u_h = 0 \quad \text{on } \partial\Omega_h, \tag{4.22b}$$

$$\alpha_\theta \widehat{M}_h = 0 \quad \text{on } \partial\Omega_h, \tag{4.22c}$$

$$\widehat{\theta}_h = 0 \quad \text{on } \partial\Omega_h, \tag{4.22d}$$

$$\widehat{T}_h, \widehat{N}_h, \text{ and } \widehat{M}_h \text{ are constants on } \partial\Omega_h. \tag{4.22e}$$

Proof. Inserting (4.21) into (4.15) we get that $\Theta_{int} + \Theta_{tr} = 0$. Since $\Theta_{int} \geq 0$, and $\Theta_{tr} \geq 0$ by (4.4), we immediately obtain (4.22a) and (4.22b). We also see that $\alpha_\theta(\widehat{M}_h - M_h) = 0$ on $\partial\Omega_h$ which implies (4.22c) since $M_h = 0$.

By (4.22a), the equation (4.1d) simplifies to $\langle \widehat{M}_h, v_4 n\rangle$ for every $v_4 \in V_h^k$. Taking $v_4 = \chi_K$, the characteristic function of the interval $K \in \Omega_h$, and varying K over all elements in Ω_h, we see that $\widehat{M}_h$ is a constant on $\partial\Omega_h$. Similarly, the simplified forms of (4.1e) and (4.1f) (since $T_h = N_h = 0$), we deduce that $\widehat{T}_h$ and $\widehat{N}_h$ are also constants on $\partial\Omega_h$. Thus, (4.22e) is proved.

To prove (4.22d), we note that (4.1c) reads $(\theta_h, v_3') + \langle \widehat{\theta}_h, v_3 n\rangle = 0$ for all $v_3 \in V_h^k$ since $M_h = 0$. Once again, setting $v_3 = \chi_K$ and varying K over all elements in Ω_h, we see that $\widehat{\theta}_h$

is constant on $\partial\Omega_h$. Since $\widehat{\theta}_h = \theta_N = 0$ by (4.1g), we readily get (4.22d). This completes the proof. □

We are now ready to prove Theorem 4.1.

Proof. (Theorem 4.1) Due to the linearity of the problem, it is enough to show that the only solution to (4.1) with data given by (4.21) is

$$T_h = N_h = M_h = \theta_h = u_h = w_h = 0 \quad \text{in } \Omega_h \tag{4.23}$$

and

$$\widehat{M_h} = \widehat{u}_h = \widehat{w}_h = 0 \quad \text{on } \mathcal{E}_h. \tag{4.24}$$

In Lemma A, we have proved that $T_h = N_h = M_h = 0$ in Ω_h. Hence it remains to prove (4.24) and that $\theta_h = u_h = w_h = 0$ in Ω_h.

We begin with proving that $\theta_h = 0$. By (4.22a), (4.22b), (4.22c), and the definition of the numerical traces, (4.2c), we have that $\widehat{\theta}_h = \theta_h$ on $\partial\Omega_h$. Thus, (4.1c) can be written as

$$-(\theta_h, v_3') + \langle \theta_h, v_3 n \rangle = 0$$

which, upon integration by parts, takes the form

$$(\theta_h', v_3) = 0 \quad \text{for all } v_3 \in V_h^k.$$

This implies that $\theta_h' = 0$ on each element $K \in \Omega_h$ and hence is a constant on each element. Since, $\widehat{\theta}_h = \theta_h$, and $\widehat{\theta}_h = 0$ on $\partial\Omega_h$ by (4.22d), we get that $\theta_h = 0$ on $\partial\Omega_h$. Since θ_h is a constant on each element we get that $\theta_h = 0$ on Ω_h.

Note that, we can now write (4.1a) and (4.1b) as

$$-(w_h, v_1') + \langle \widehat{w}_h, v_1 n \rangle + (\kappa u_h, v_1) = 0,$$

$$-(u_h, v_2') + \langle \widehat{u}_h, v_2 n \rangle - (\kappa w_h, v_2) = 0,$$

for all v_1, $v_2 \in V_h^k$. Integrating by parts on both of the first terms on the left-hand side and noting that we have (4.22b) on $\partial\Omega_h$, we get that

$$(w_h', v_1) + (\kappa u_h, v_1) = 0,$$

$$(u_h', v_2) - (\kappa w_h, v_2) = 0.$$

Using the assumption (4.3), we can now prove that $u_h = w_h = 0$ in Ω_h. For details, see Appendix A in [42]. This completes the proof of (4.23). Consequently, by (4.22b), we get that $\widehat{u}_h = \widehat{w}_h = 0$ on $\partial\Omega_h$.

It remains to prove that $\widehat{M}_h = 0$ on $\partial\Omega_h$. By (4.22e), $\widehat{M}_h$ is a constant on $\partial\Omega_h$, and by (4.22a), the equation (4.1d) takes the form

$$\langle \widehat{M}_h, v_4 n \rangle = 0 \quad \text{for all } v_4 \in V_h^k.$$

Now, for $k \geq 1$, let $K = (a, b)$ be and arbitrary element and let v_4 be the linear function on K such that $v_4(a) = 0$, $v_4(b) = 1$, and v_4 is zero on all other elements. Then the above equation implies that $\widehat{M}_h(b^-) = 0$. Thus, since $\widehat{M}_h$ is a constant on $\partial\Omega_h$, we see that $\widehat{M}_h = 0$ on $\partial\Omega_h$. For $k = 0$, however, we can not use the linear test function v_4 above. On the other hand, the assumption (4.5) together with (4.22d) implies that $\widehat{M}_h = 0$ on at least one node of the mesh. But since $\widehat{M}_h$ is a constant, it must be zero on all of $\partial\Omega_h$. This completes the proof of (4.24) and that of the theorem. □

4.6.2 Approximation properties of the projection: Proof of Theorem 4.3

In this subsection, we provide a detailed proof of Theorem 4.3. We only give the proof for $k \geq 1$. The proof for $k = 0$ is similar and easier.

Fix an interval $K = (x_L, x_R) \in \Omega_h$ and set

$$d_z := z_k - \Pi_z z, \qquad \boldsymbol{d} := (d_T, d_N, d_M, d_\theta, d_u, d_w),$$
$$g_z := z - z_k \qquad\quad \boldsymbol{g} := (g_T, g_N, g_M, g_\theta, g_u, g_w),$$

where z_k denotes the L^2-projection of z into $\mathcal{P}_k(K)$. Since $\boldsymbol{\delta} = \boldsymbol{g} + \boldsymbol{d}$, we only need to estimate $\boldsymbol{d}$. To do that, we proceed as follows. From the definition of the projection (4.7) and the definition of the L^2-projection into $\mathcal{P}_k(K)$, we have

$$\begin{aligned} (d_T, v_1)_K = (d_N, v_2)_K = (d_M, v_3)_K = 0, \\ (d_\theta, v_4)_K = (d_u, v_5)_K = (d_w, v_6)_K = 0, \end{aligned} \tag{4.25}$$

for all $(v_1, \ldots, v_6) \in \boldsymbol{\mathcal{P}}^{k-1}(K)$, and

$$\begin{bmatrix} d_\theta \\ d_N \\ d_T \end{bmatrix} n - \mathtt{S} \begin{bmatrix} d_M \\ d_u \\ d_w \end{bmatrix} = \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix} n - \mathtt{S} \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix} \qquad \text{on } \partial K. \tag{4.26}$$

By equations (4.25), we see that we can write $d_z = C_z\, L_k$ where L_k denotes the scaled Legendre polynomial of degree k. Hence, evaluating (4.26) at the left end of the interval K and noting that $L_k(x_L) = (-1)^k$ and $n(x_L) = -1$ we get

$$\begin{bmatrix} C_\theta \\ C_N \\ C_T \end{bmatrix} + \mathtt{S}^+ \begin{bmatrix} C_M \\ C_u \\ C_w \end{bmatrix} = (-1)^k \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^+ + (-1)^k \mathtt{S}^+ \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^+ .$$

Similarly, evaluating (4.26) at the right end of the interval K and noting that $L_k(x_R) = 1$ and $n(x_R) = 1$ we get

$$\begin{bmatrix} C_\theta \\ C_N \\ C_T \end{bmatrix} - \mathtt{S}^- \begin{bmatrix} C_M \\ C_u \\ C_w \end{bmatrix} = \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^- - \mathtt{S}^- \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^- .$$

Consequently, we can write (4.26) in the following block-matrix form

$$\begin{bmatrix} I & \mathtt{S}^+ \\ I & -\mathtt{S}^- \end{bmatrix} \begin{bmatrix} \begin{bmatrix} C_\theta \\ C_N \\ C_T \end{bmatrix} \\ \begin{bmatrix} C_M \\ C_u \\ C_w \end{bmatrix} \end{bmatrix} = \begin{bmatrix} (-1)^k \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^+ + (-1)^k \mathtt{S}^+ \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^+ \\ \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^- - \mathtt{S}^- \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^- \end{bmatrix}$$

where I denotes the 3×3 identity matrix. It is now evident that the system has a unique solution if and only if the matrix $(\mathtt{S}^- + \mathtt{S}^+)$ is non-singular. Assuming that this is the case we obtain, by elementary block-row elimination and back-substitution and some algebraic manipulation, that

$$\begin{bmatrix} C_M \\ C_u \\ C_w \end{bmatrix} = -\mathtt{P} \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^+ - \mathtt{P}\mathtt{S}^+ \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^+ + (-1)^k \mathtt{P} \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^- + (-1)^k \mathtt{P}\mathtt{S}^- \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^-$$

and

$$\begin{bmatrix} C_\theta \\ C_N \\ C_T \end{bmatrix} = (-1)^k \mathtt{S}^- \mathtt{P} \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^+ + (-1)^k \mathtt{S}^- \mathtt{P} \mathtt{S}^+ \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^+ + \mathtt{S}^+ \mathtt{P} \begin{bmatrix} g_\theta \\ g_N \\ g_T \end{bmatrix}^- - \mathtt{S}^+ \mathtt{P} \begin{bmatrix} g_M \\ g_u \\ g_w \end{bmatrix}^- .$$

Thus, we conclude that

$$\begin{aligned} \|\boldsymbol{d}\|_K &= \|L_k\|_K \left(|C_T| + |C_N| + |C_M| + |C_\theta| + |C_u| + |C_w|\right) \\ &\leq C_{\mathtt{s}} \|L_k\|_K \|\boldsymbol{g}\|_{\partial K} \\ &\leq C_{\mathtt{s}} h^{1/2} \|\boldsymbol{g}\|_{\partial K} \\ &\leq C C_{\mathtt{s}} \|\boldsymbol{g}\|_K \\ &\leq C C_{\mathtt{s}}\, h^s |\mathbf{z}|_{s,K}, \end{aligned}$$

for all $1 \leq s \leq k+1$, by the trace inequality and the approximation properties of the L^2-projection.

By triangle inequality, we have

$$\|\boldsymbol{\delta}\|_K \leq \|\boldsymbol{d}\|_K + \|\boldsymbol{g}\|_K ,$$

and the estimate of Theorem 4.3 readily follows by adding over all elements $K \in \Omega_h$. This completes the proof.

4.6.3 Estimates of the projection of the error: Proof of Theorem 4.4.

This subsection is devoted to the proof of Theorem 4.4. We proceed in two steps in the first of which, we use a key identity obtained by duality to prove Theorem 4.4. In the second step, we prove the identity.

Step 1: The duality identity and the proof of Theorem 4.4 Our proof will is based on the following auxiliary result.

Lemma 4.9. *For any* $(\eta_T, \eta_N, \eta_M, \eta_\theta, \eta_u, \eta_w) \in \boldsymbol{L}^2(\Omega_h)$*, set*

$$\mathcal{E}_z := (\Pi_z e_z, \eta_z) \quad \textit{and} \quad \mathcal{E} = \mathcal{E}_T + \mathcal{E}_N + \mathcal{E}_M + \mathcal{E}_\theta + \mathcal{E}_u + \mathcal{E}_w.$$

Then

$$\begin{aligned}
\mathcal{E} = \quad & (\Pi_\theta e_\theta, \delta_{\psi_T}) \quad + (\Pi_M e_M, \delta_{\psi_M}) \; - (\Pi_T e_T, \delta_{\psi_\theta}) \\
& - (\delta_\theta, \Pi_T \psi_T) \quad - (\delta_M, \Pi_M \psi_M) \; + (\delta_T, \Pi_\theta \psi_\theta) \\
& - d^2(\Pi_N e_N, \delta_{\psi_N}) - d^2(\Pi_T e_T, \delta_{\psi_T}) + d^2(\delta_N, \Pi_N \psi_N) + d^2(\delta_T, \Pi_T \psi_T) \\
& - (\Pi_w e_w, \kappa \delta_{\psi_N}) \; + (\Pi_u e_u, \kappa \delta_{\psi_T}) \; - (\Pi_N e_N, \kappa \delta_{\psi_w}) \; + (\Pi_T e_T, \kappa \delta_{\psi_u}) \\
& + (\kappa \delta_w, \Pi_N \psi_N) \; - (\kappa \delta_u, \Pi_T \psi_T) \; + (\kappa \delta_N, \Pi_w \psi_w) \; - (\kappa \delta_T, \Pi_u \psi_u).
\end{aligned}$$

Here, on each $K \in \Omega_h$*, we take* S^t *as the stabilization function for defining the projection* $\boldsymbol{\Pi\psi}$*.*

We delay the proof of this identity to the end of this subsection. We are now ready to prove Theorem 4.4.

Proof. (Theorem 4.4) We first consider the case $k \geq 1$. Setting

$$\boldsymbol{\eta} = (\eta_T, \eta_N, \eta_M, \eta_\theta, \eta_u, \eta_w) = (\Pi_T e_T, \Pi_N e_N, \Pi_M e_M, \Pi_\theta e_\theta, \Pi_u e_u, \Pi_w e_w) = \boldsymbol{\Pi e}$$

in the identity of Lemma 4.9 gives

$$\begin{aligned}
\|\mathbf{\Pi e}\|^2 = \quad & (\Pi_\theta e_\theta, \delta_{\psi_T}) \quad + (\Pi_M e_M, \delta_{\psi_M}) \; - (\Pi_T e_T, \delta_{\psi_\theta}) \\
& - (\delta_\theta, \Pi_T \psi_T) \quad - (\delta_M, \Pi_M \psi_M) \; + (\delta_T, \Pi_\theta \psi_\theta) \\
& - d^2(\Pi_N e_N, \delta_{\psi_N}) - d^2(\Pi_T e_T, \delta_{\psi_T}) + d^2(\delta_N, \Pi_N \psi_N) + d^2(\delta_T, \Pi_T \psi_T) \\
& - (\Pi_w e_w, \kappa \delta_{\psi_N}) \; + (\Pi_u e_u, \kappa \delta_{\psi_T}) \; - (\Pi_N e_N, \kappa \delta_{\psi_w}) \; + (\Pi_T e_T, \kappa \delta_{\psi_u}) \\
& + (\kappa \delta_w, \Pi_N \psi_N) \; - (\kappa \delta_u, \Pi_T \psi_T) \; + (\kappa \delta_N, \Pi_w \psi_w) \; - (\kappa \delta_T, \Pi_u \psi_u).
\end{aligned}$$

Using the fact that $\Pi_z \psi_z = \psi_z - \delta_{\psi_z}$, we get

$$\|\mathbf{\Pi e}\|^2 = T_1 + T_2 + \cdots + T_9$$

where

$$\begin{aligned}
T_1 &= (\delta_T, \psi_\theta) \quad - (\delta_T, \delta_{\psi_\theta}) \quad - (\Pi_T e_T, \delta_{\psi_\theta}), \\
T_2 &= d^2(\delta_T, \psi_T) \; - d^2(\delta_T, \delta_{\psi_T}) \; - d^2(\Pi_T e_T, \delta_{\psi_T}), \\
T_3 &= d^2(\delta_N, \psi_N) - d^2(\delta_N, \delta_{\psi_N}) - d^2(\Pi_N e_N, \delta_{\psi_N}),
\end{aligned} \tag{4.27}$$

$$\begin{aligned}
T_4 &= -(\delta_M, \psi_M) + (\delta_M, \delta_{\psi_M}) + (\Pi_M e_M, \delta_{\psi_M}), \\
T_5 &= -(\delta_\theta, \psi_T) \; + (\delta_\theta, \delta_{\psi_T}) \; + (\Pi_\theta e_\theta, \delta_{\psi_T}), \\
T_6 &= -(\delta_T, \kappa \psi_u) + (\delta_T, \kappa \delta_{\psi_u}) + (\Pi_T e_T, \kappa \delta_{\psi_u}),
\end{aligned} \tag{4.28}$$

$$\begin{aligned}
T_7 &= (\delta_N, \kappa \psi_w) \; - (\delta_N, \kappa \delta_{\psi_w}) - (\Pi_N e_N, \kappa \delta_{\psi_w}), \\
T_8 &= -(\delta_u, \kappa \psi_T) + (\delta_u, \kappa \delta_{\psi_T}) \; + (\Pi_u e_u, \kappa \delta_{\psi_T}), \\
T_9 &= (\delta_w, \kappa \psi_N) \; - (\delta_w, \kappa \delta_{\psi_N}) - (\Pi_w e_w, \kappa \delta_{\psi_N}),
\end{aligned} \tag{4.29}$$

By the orthogonality property of the projection, (4.7), we can rewrite these equations as

$$\begin{aligned} T_1 &= (\delta_T, \psi_\theta - (\psi_\theta)_{k-1}) \quad - (\delta_T, \delta_{\psi_\theta}) \quad - (\Pi_T e_T, \delta_{\psi_\theta}), \\ T_2 &= d^2(\delta_T, \psi_T - (\psi_T)_{k-1}) \; - d^2(\delta_T, \delta_{\psi_T}) \; - d^2(\Pi_T e_T, \delta_{\psi_T}), \\ T_3 &= d^2(\delta_N, \psi_N - (\psi_N)_{k-1}) - d^2(\delta_N, \delta_{\psi_N}) - d^2(\Pi_N e_N, \delta_{\psi_N}), \end{aligned} \tag{4.30}$$

$$\begin{aligned} T_4 &= -(\delta_M, \psi_M - (\psi_M)_{k-1}) \; + (\delta_M, \delta_{\psi_M}) + (\Pi_M e_M, \delta_{\psi_M}), \\ T_5 &= -(\delta_\theta, \psi_T - (\psi_T)_{k-1}) \quad + (\delta_\theta, \delta_{\psi_T}) \quad + (\Pi_\theta e_\theta, \delta_{\psi_T}), \\ T_6 &= -(\delta_T, \kappa\psi_u - (\kappa\psi_u)_{k-1}) + (\delta_T, \kappa\delta_{\psi_u}) + (\Pi_T e_T, \kappa\delta_{\psi_u}), \end{aligned} \tag{4.31}$$

$$\begin{aligned} T_7 &= (\delta_N, \kappa\psi_w - (\kappa\psi_w)_{k-1}) \; - (\delta_N, \kappa\delta_{\psi_w}) - (\Pi_N e_N, \kappa\delta_{\psi_w}), \\ T_8 &= -(\delta_u, \kappa\psi_T - (\kappa\psi_T)_{k-1}) + (\delta_u, \kappa\delta_{\psi_T}) \; + (\Pi_u e_u, \kappa\delta_{\psi_T}), \\ T_9 &= (\delta_w, \kappa\psi_N - (\kappa\psi_N)_{k-1}) \; - (\delta_w, \kappa\delta_{\psi_N}) - (\Pi_w e_w, \kappa\delta_{\psi_N}). \end{aligned} \tag{4.32}$$

An estimate on $\|\mathbf{\Pi e}\|$ now follows by estimating T_i for $i = 1, \ldots, 9$. We only show the details of how to estimate T_6, since the remaining terms can be estimated in a similar fashion. Applying the Cauchy-Schwarz inequality to each term in T_6, we get

$$|T_6| \leq \|\delta_T\| \, \|\kappa\psi_u - (\kappa\psi_u)_{k-1}\| + (\|\delta_T\| + \|\Pi_T e_T\|) \, \|\kappa\delta_{\psi_u}\| \, .$$

By the approximation properties of the L^2-projection, we get that

$$\begin{aligned} |T_6| &\leq Ch \, \|\delta_T\| \, \|\kappa\psi_u\|_1 + \|\kappa\|_\infty (\|\delta_T\| + \|\Pi_T e_T\|) \, \|\delta_{\psi_u}\| \\ &\leq Ch \, \|\delta_T\| \, \|\kappa\|_1 \, \|\psi_u\|_1 + \|\kappa\|_\infty (\|\delta_T\| + \|\Pi_T e_T\|) \, \|\delta_{\psi_u}\| \, . \end{aligned}$$

By the assumption that κ is very smooth, and by Theorem 4.3, we get that

$$\begin{aligned} |T_6| &\leq Ch \, \|\delta_T\| \, \|\psi_u\|_1 + Ch(\|\delta_T\| + \|\Pi_T e_T\|)|\boldsymbol{\psi}|_1 \\ &\leq Ch \, \|\boldsymbol{\delta}\| \, \|\boldsymbol{\psi}\|_1 + Ch(\|\boldsymbol{\delta}\| + \|\mathbf{\Pi e}\|) \, \|\boldsymbol{\psi}\|_1 \, . \end{aligned}$$

By the elliptic regularity estimate (4.11), we have

$$\|\boldsymbol{\psi}\|_1 \leq C_{reg} \|\boldsymbol{\Pi e}\|$$

and hence

$$|T_6| \leq CC_{reg}h \|\boldsymbol{\delta}\| \|\boldsymbol{\Pi e}\| + CC_{reg}h \|\boldsymbol{\delta}\| \|\boldsymbol{\Pi e}\|^2 .$$

Estimating the remaining terms similarly we obtain

$$\begin{aligned} \|\boldsymbol{\Pi e}\|^2 &\leq |T_1| + |T_2| + \cdots + |T_9| \\ &\leq CC_{reg}h \|\boldsymbol{\delta}\| \|\boldsymbol{\Pi e}\| + CC_{reg}h \|\boldsymbol{\delta}\| \|\boldsymbol{\Pi e}\|^2 . \end{aligned}$$

If we assume that h is small enough so that $CC_{reg}h < 1$ then

$$\|\boldsymbol{\Pi e}\|^2 \leq C\, C_{reg}h \|\boldsymbol{\delta}\| \|\boldsymbol{\Pi e}\| ,$$

and the first estimate of Theorem (4.4) follows.

Next we consider the case $k = 0$. In this case (4.27)-(4.29) are still valid, but we do not have (4.30)-(4.32) since the L^2-projection into polynomials of degree $k-1$ is no longer defined. Nevertheless, we can still estimate T_i for $i = 1, \ldots, 9$ in their form given by (4.27)-(4.29). We provide the details for only T_2. Applying Cauchy-Schwarz inequality to each term in T_2 we get

$$\begin{aligned} |T_2| &\leq d^2 \|\delta_T\| \|\psi_T\| + d^2(\|\delta_T\| + \|\Pi_T e_T\|) \|\delta_{\psi_T}\| \\ &\leq \|\delta_T\| \|\psi_T\| + (\|\delta_T\| + \|\Pi_T e_T\|) \|\delta_{\psi_T}\| \end{aligned}$$

since $0 < d < 1$. By Theorem 4.3 we have that

$$\begin{aligned} |T_2| &\leq \|\delta_T\| \|\psi_T\| + (\|\delta_T\| + \|\Pi_T e_T\|) Ch \|\boldsymbol{\psi}\|_1 \\ &\leq \|\boldsymbol{\delta}\| \|\boldsymbol{\psi}\| + (\|\boldsymbol{\delta}\| + \|\boldsymbol{\Pi e}\|) Ch \|\boldsymbol{\psi}\|_1 , \end{aligned}$$

and, by the elliptic regularity estimate (4.11) we have

$$|T_2| \leq CC_{reg} \|\boldsymbol{\delta}\| \|\boldsymbol{\Pi e}\| + CC_{reg} h \|\boldsymbol{\Pi e}\|^2 .$$

Since the remaining terms can be estimated in a similar fashion, we obtain

$$\|\boldsymbol{\Pi e}\|^2 \leq CC_{reg} \|\boldsymbol{\delta}\| \|\boldsymbol{\Pi e}\| + CC_{reg} h \|\boldsymbol{\Pi e}\|^2 .$$

The second estimate of Theorem (4.4) now follows if we assume that $CC_{reg}h < 1$. This completes the proof. □

Step 2: Proof of the duality identity of Lemma 4.9. To prove Lemma 4.9, we begin by obtaining a couple of auxiliary identities. The first is the following.

Lemma 4.10. *Let $\boldsymbol{v} = (v_1, \ldots, v_6) \in \boldsymbol{H}^1(\Omega_h)$ and we take $\mathtt{S}^t$ as the stabilization function of the projection $\boldsymbol{\Pi v}$. Then*

$$- \langle \widehat{e}_T - e_T, \delta_{v_6} n\rangle - \langle \widehat{e}_N - e_N, \delta_{v_5} n\rangle + \langle \widehat{e}_M - e_M, \delta_{v_4} n\rangle$$

$$- \langle \widehat{e}_\theta - e_\theta, \delta_{v_3} n\rangle \ + \langle \widehat{e}_u - e_u, \delta_{v_2} n\rangle \ + \langle \widehat{e}_w - e_w, \delta_{v_1} n\rangle = 0.$$

Proof. Let Θ be the left-hand side of the identity we want to prove, that is,

$$\Theta := -\langle \begin{bmatrix} \widehat{e}_\theta - e_\theta \\ \widehat{e}_N - e_N \\ \widehat{e}_T - e_T \end{bmatrix}, \begin{bmatrix} \delta_{v_3} \\ \delta_{v_5} \\ \delta_{v_6} \end{bmatrix} n\rangle + \langle \begin{bmatrix} \widehat{e}_M - e_M \\ \widehat{e}_u - e_u \\ \widehat{e}_w - e_w \end{bmatrix}, \begin{bmatrix} \delta_{v_4} \\ \delta_{v_2} \\ \delta_{v_1} \end{bmatrix} n\rangle$$

with the obvious extension of the definition of $\langle \cdot, \cdot \rangle$ for vector-valued functions. Noting that

$$\widehat{e}_z - e_z = z_h - \widehat{z}_h,$$

and that, by the definition of the numerical traces, (4.2c), we have

$$\begin{bmatrix} \widehat{e}_\theta - e_\theta \\ \widehat{e}_N - e_N \\ \widehat{e}_T - e_T \end{bmatrix} = \begin{bmatrix} \theta_h - \widehat{\theta}_h \\ N_h - \widehat{N}_h \\ T_h - \widehat{T}_h \end{bmatrix} = \mathtt{S} \begin{bmatrix} M_h - \widehat{M}_h \\ u_h - \widehat{u}_h \\ w_h - \widehat{w}_h \end{bmatrix} n,$$

we get

$$\Theta = -\langle \mathtt{S} \begin{bmatrix} M_h - \widehat{M}_h \\ u_h - \widehat{u}_h \\ w_h - \widehat{w}_h \end{bmatrix}, \begin{bmatrix} \delta_{v_3} \\ \delta_{v_5} \\ \delta_{v_6} \end{bmatrix} \rangle + \langle \begin{bmatrix} M_h - \widehat{M}_h \\ u_h - \widehat{u}_h \\ w_h - \widehat{w}_h \end{bmatrix}, \mathtt{S}^t \begin{bmatrix} \delta_{v_3} \\ \delta_{v_5} \\ \delta_{v_6} \end{bmatrix} \rangle = 0$$

because

$$\begin{bmatrix} \delta_{v_4} \\ \delta_{v_2} \\ \delta_{v_1} \end{bmatrix} = \mathtt{S}^t \begin{bmatrix} \delta_{v_3} \\ \delta_{v_5} \\ \delta_{v_6} \end{bmatrix} n$$

by (4.7c). This completes the proof. □

Lemma 4.11. *Let u_i, $v_i \in H^1(\Omega_h)$ for $i = 1, \ldots, 6$, with the stabilization functions* $\mathtt{S}$ *and* $\mathtt{S}^t$*, respectively. Then*

$$\langle \delta_{u_6}, \delta_{v_1} n \rangle + \langle \delta_{u_5}, \delta_{v_2} n \rangle - \langle \delta_{u_4}, \delta_{v_3} n \rangle$$

$$+ \langle \delta_{u_3}, \delta_{v_4} n \rangle - \langle \delta_{u_2}, \delta_{v_5} n \rangle - \langle \delta_{u_1}, \delta_{v_6} n \rangle = 0.$$

Proof. Let Θ be the left-hand side of the identity we want to prove, that is,

$$\Theta := -\langle \begin{bmatrix} \delta_{u_4} \\ \delta_{u_2} \\ \delta_{u_1} \end{bmatrix}, \begin{bmatrix} \delta_{v_3} \\ \delta_{v_5} \\ \delta_{v_6} \end{bmatrix} n \rangle + \langle \begin{bmatrix} \delta_{u_6} \\ \delta_{u_5} \\ \delta_{u_3} \end{bmatrix}, \begin{bmatrix} \delta_{v_1} \\ \delta_{v_2} \\ \delta_{v_4} \end{bmatrix} n \rangle.$$

Since, by (4.7c), we have that

$$\begin{bmatrix} \delta_{u_4} \\ \delta_{u_2} \\ \delta_{u_1} \end{bmatrix} = \mathtt{S} \begin{bmatrix} \delta_{u_3} \\ \delta_{u_5} \\ \delta_{u_6} \end{bmatrix} n \quad \text{and} \quad \begin{bmatrix} \delta_{v_1} \\ \delta_{v_2} \\ \delta_{v_4} \end{bmatrix} = \mathtt{S}^t \begin{bmatrix} \delta_{v_6} \\ \delta_{v_5} \\ \delta_{v_3} \end{bmatrix} n,$$

we readily obtain that

$$\Theta = -\langle \mathtt{S} \begin{bmatrix} \delta_{u_3} \\ \delta_{u_5} \\ \delta_{u_6} \end{bmatrix} n, \begin{bmatrix} \delta_{v_3} \\ \delta_{v_5} \\ \delta_{v_6} \end{bmatrix} n \rangle + \langle \begin{bmatrix} \delta_{u_6} \\ \delta_{u_5} \\ \delta_{u_3} \end{bmatrix}, \mathtt{S}^t \begin{bmatrix} \delta_{v_6} \\ \delta_{v_5} \\ \delta_{v_3} \end{bmatrix} \rangle = 0.$$

This completes the proof. □

We are now ready to prove Lemma 4.9.

Proof. (Lemma 4.9) By the definition of $\mathcal{E}$ and the equations defining the dual solution (4.10), we have

$$\begin{aligned}
\mathcal{E} = \quad & (\Pi_T e_T, \psi_w') \; - (\Pi_T e_T, \psi_\theta) \quad + (\Pi_T e_T, \kappa\psi_u) \; - d^2(\Pi_T e_T, \psi_T) \\
& + (\Pi_N e_N, \psi_u') \qquad\qquad\qquad - (\Pi_N e_N, \kappa\psi_w) - d^2(\Pi_N e_N, \psi_N) \\
& - (\Pi_M e_M, \psi_\theta') + (\Pi_M e_M, \psi_M) + (\Pi_\theta e_\theta, \psi_M') \quad + (\Pi_\theta e_\theta, \psi_T) \\
& - (\Pi_u e_u, \psi_N') \; + (\Pi_u e_u, \kappa\psi_T) \; - (\Pi_w e_w, \psi_T') \quad - (\Pi_w e_w, \kappa\psi_N).
\end{aligned}$$

Since, for any pair, (e_z, ψ_v), we have

$$\begin{aligned}
(\Pi_z e_z, \psi_v') &= (\Pi_z e_z, (\Pi_v \psi_v)') + (\Pi_z e_z, \delta_{\psi_v}') \\
&= (\Pi_z e_z, (\Pi_v \psi_v)') + \langle \Pi_z e_z, \delta_{\psi_v} n\rangle - ((\Pi_z e_z)', \delta_{\psi_v}) \\
&= (\Pi_z e_z, (\Pi_v \psi_v)') + \langle \Pi_z e_z, \delta_{\psi_v} n\rangle,
\end{aligned}$$

by the orthogonality properties (4.7a)-(4.7b) of the projection. Hence

$$\begin{aligned}
\mathcal{E} = \quad & (\Pi_T e_T, (\Pi_w \psi_w)') \; - (\Pi_T e_T, \psi_\theta) \quad + (\Pi_T e_T, \kappa\psi_u) \qquad - d^2(\Pi_T e_T, \psi_T) \\
& + (\Pi_N e_N, (\Pi_u \psi_u)') \qquad\qquad\qquad - (\Pi_N e_N, \kappa\psi_w) \qquad - d^2(\Pi_N e_N, \psi_N) \\
& - (\Pi_M e_M, (\Pi_\theta \psi_\theta)') + (\Pi_M e_M, \psi_M) \; + (\Pi_\theta e_\theta, (\Pi_M \psi_M)') + (\Pi_\theta e_\theta, \psi_T) \\
& - (\Pi_u e_u, (\Pi_N \psi_N)') + (\Pi_u e_u, \kappa\psi_T) \quad - (\Pi_w e_w, (\Pi_T \psi_T)') - (\Pi_w e_w, \kappa\psi_N) \\
& + \langle \Pi_T e_T, \delta_{\psi_w} n\rangle \qquad + \langle \Pi_N e_N, \delta_{\psi_u} n\rangle - \langle \Pi_M e_M, \delta_{\psi_\theta} n\rangle \\
& + \langle \Pi_\theta e_\theta, \delta_{\psi_M} n\rangle \qquad - \langle \Pi_u e_u, \delta_{\psi_N} n\rangle \; - \langle \Pi_w e_w, \delta_{\psi_T} n\rangle.
\end{aligned}$$

Taking

$$(v_1, v_2, v_3, v_4, v_5, v_6) = (-\Pi_T\psi_T, -\Pi_N\psi_N, \Pi_M\psi_M, -\Pi_\theta\psi_\theta, \Pi_u\psi_u, \Pi_w\psi_w)$$

in the error equations (4.8) and carrying out some very simple algebraic manipulations, we obtain

$$\begin{aligned}
\mathcal{E} = H &+ (\Pi_\theta e_\theta, \delta_{\psi_T}) \quad + (\Pi_M e_M, \delta_{\psi_M}) \; - (\Pi_T e_T, \delta_{\psi_\theta}) \\
&- (\delta_\theta, \Pi_T\psi_T) \quad - (\delta_M, \Pi_M\psi_M) \; + (\delta_T, \Pi_\theta\psi_\theta) \\
&- d^2(\Pi_N e_N, \delta_{\psi_N}) - d^2(\Pi_T e_T, \delta_{\psi_T}) + d^2(\delta_N, \Pi_N\psi_N) + d^2(\delta_T, \Pi_T\psi_T) \\
&- (\Pi_w e_w, \kappa\delta_{\psi_N}) \; + (\Pi_u e_u, \kappa\delta_{\psi_T}) \; - (\Pi_N e_N, \kappa\delta_{\psi_w}) \; + (\Pi_T e_T, \kappa\delta_{\psi_u}) \\
&+ (\kappa\delta_w, \Pi_N\psi_N) \; - (\kappa\delta_u, \Pi_T\psi_T) \; + (\kappa\delta_N, \Pi_w\psi_w) \; - (\kappa\delta_T, \Pi_u\psi_u)
\end{aligned}$$

where

$$\begin{aligned}
H = \; &\langle \widehat{e}_T, \Pi_w\psi_w n\rangle \; + \langle \Pi_T e_T, \delta_{\psi_w} n\rangle \; + \langle \widehat{e}_N, \Pi_u\psi_u n\rangle \; + \langle \Pi_N e_N, \delta_{\psi_u} n\rangle \\
&- \langle \widehat{e}_M, \Pi_\theta\psi_\theta n\rangle \; - \langle \Pi_M e_M, \delta_{\psi_\theta} n\rangle + \langle \widehat{e}_\theta, \Pi_M\psi_M n\rangle + \langle \Pi_\theta e_\theta, \delta_{\psi_M} n\rangle \\
&- \langle \widehat{e}_u, \Pi_N\psi_N n\rangle - \langle \Pi_u e_u, \delta_{\psi_N} n\rangle \; - \langle \widehat{e}_w, \Pi_T\psi_T n\rangle \; - \langle \Pi_w e_w, \delta_{\psi_T} n\rangle.
\end{aligned}$$

It remains to show that $H = 0$.

Since ψ_M, ψ_u, and ψ_w are single-valued on $\mathcal{E}_h$, and $\psi_u = \psi_w = 0$ on $\partial\Omega$, we can take $\mathsf{m} = \psi_M$, $\mathsf{u} = \psi_u$, and $\mathsf{w} = \psi_w$ in the error equations (4.8g)-(4.8i), respectively to get

$$\langle \widehat{e}_\theta, \psi_M n\rangle = \langle \widehat{e}_N, \psi_u n\rangle = \langle \widehat{e}_T, \psi_w n\rangle = 0.$$

Moreover, since $\widehat{e}_M$, $\widehat{e}_u$, and $\widehat{e}_w$ are single valued on $\mathcal{E}_h$, and $\widehat{e}_u = 0$, $\widehat{e}_w = 0$, and $\psi_\theta = 0$ on $\partial\Omega$, we have

$$\langle \widehat{e}_M, \psi_\theta n\rangle = \langle \widehat{e}_u, \psi_N n\rangle = \langle \widehat{e}_w, \psi_T n\rangle = 0.$$

This implies that

$$\begin{aligned} H = \quad & \langle \widehat{e}_T, (\Pi_w \psi_w - \psi_w) n \rangle \quad + \langle \Pi_T e_T, \delta_{\psi_w} n \rangle \\ & + \langle \widehat{e}_N, (\Pi_u \psi_u - \psi_u) n \rangle \quad + \langle \Pi_N e_N, \delta_{\psi_u} n \rangle \\ & - \langle \widehat{e}_M, (\Pi_\theta \psi_\theta - \psi_\theta) n \rangle \quad - \langle \Pi_M e_M, \delta_{\psi_\theta} n \rangle \\ & + \langle \widehat{e}_\theta, (\Pi_M \psi_M - \psi_M) n \rangle + \langle \Pi_\theta e_\theta, \delta_{\psi_M} n \rangle \\ & - \langle \widehat{e}_u, (\Pi_N \psi_N - \psi_N) n \rangle \ - \langle \Pi_u e_u, \delta_{\psi_N} n \rangle \\ & - \langle \widehat{e}_w, (\Pi_T \psi_T - \psi_T) n \rangle \quad - \langle \Pi_w e_w, \delta_{\psi_T} n \rangle \\ = \quad & H_1 + H_2 \end{aligned}$$

where

$$\begin{aligned} H_1 = & - \langle \widehat{e}_T - e_T, \delta_{\psi_w} n \rangle - \langle \widehat{e}_N - e_N, \delta_{\psi_u} n \rangle + \langle \widehat{e}_M - e_M, \delta_{\psi_\theta} n \rangle \\ & - \langle \widehat{e}_\theta - e_\theta, \delta_{\psi_M} n \rangle + \langle \widehat{e}_u - e_u, \delta_{\psi_N} n \rangle + \langle \widehat{e}_w - e_w, \delta_{\psi_T} n \rangle, \end{aligned}$$

and

$$\begin{aligned} H_2 = & - \langle \delta_T, \delta_{\psi_w} n \rangle - \langle \delta_N, \delta_{\psi_u} n \rangle + \langle \delta_M, \delta_{\psi_\theta} n \rangle \\ & - \langle \delta_\theta, \delta_{\psi_M} n \rangle + \langle \delta_u, \delta_{\psi_N} n \rangle + \langle \delta_w, \delta_{\psi_T} n \rangle. \end{aligned}$$

But $H_1 = 0$ by Lemma 4.10 with $(v_1, \ldots, v_6) = (\psi_T, \psi_N, \psi_M, \psi_\theta, \psi_u, \psi_w)$, and $H_2 = 0$ by Lemma 4.11 with $(u_1, \ldots, u_6) = (T, N, M, \theta, u, w)$ and $(v_1, \ldots, v_6) = (\psi_T, \psi_N, \psi_M, \psi_\theta, \psi_u, \psi_w)$. This completes the proof. □

4.6.4 Nodal superconvergence: Proof of Theorem 4.6

To prove this theorem we proceed in two steps. In the first, we obtain representation formulas for the errors in the numerical traces. In the second, we use approximation results to estimate

them. We prove the result only for $k \geq 1$; the proof for the case $k = 0$ is not difficult.

To simplify notation, we fix an arbitrary node $x_i \in \mathcal{E}_h$, and an arbitrary unknown $z \in \{T, N, M, \theta, u, w\}$ and drop the superindex and the second subindex from the Green's functions defined by (4.12), (4.13), and (4.14), more explicitly, we will write

$$(G_T, G_N, G_M, G_\theta, G_u, G_w)$$

instead of

$$(G^z_{T,x_i}, G^z_{N,x_i}, G^z_{M,x_i}, G^z_{\theta,x_i}, G^z_{u,x_i}, G^z_{w,x_i}).$$

Step 1: Representation of the errors The following lemma provides a representation formula for the errors in the numerical traces.

Lemma 4.12. *We have that* $\widehat{e}_z(x_i) = \Gamma_1 + \Gamma_2$ *where*

$$\begin{aligned}\Gamma_1 = \quad & (w' - (w')_{k-1}, \delta_{G_T}) \quad + (u' - (u')_{k-1}, \delta_{G_N}) \quad - (\theta' - (\theta')_{k-1}, \delta_{G_M}) \\ & + (M' - (M')_{k-1}, \delta_{G_\theta}) - (N' - (N')_{k-1}, \delta_{G_u}) - (T' - (T')_{k-1}, \delta_{G_w})\end{aligned}$$

and

$$\begin{aligned}\Gamma_2 = \quad & (e_\theta + \kappa e_u - d^2 e_T, \delta_{G_T}) - (\kappa e_w + d^2 e_N, \delta_{G_N}) + (e_M, \delta_{G_M}) \\ & - (e_T, \delta_{G_\theta}) \qquad\qquad + (\kappa e_T, \delta_{G_u}) \qquad\qquad - (\kappa e_N, \delta_{G_w}).\end{aligned}$$

To prove this lemma we need an auxiliary result which establishes a relation between the errors in the numerical traces and the Green's functions.

Lemma 4.13. *Set*

$$\begin{aligned}\Theta := \quad & \langle \widehat{e}_w, G_T\, n\rangle \; + \langle \widehat{e}_u, G_N\, n\rangle - \langle \widehat{e}_\theta, G_M\, n\rangle \\ & + \langle \widehat{e}_M, G_\theta\, n\rangle - \langle \widehat{e}_N, G_u\, n\rangle - \langle \widehat{e}_T, G_w\, n\rangle\end{aligned}$$

Then, we have $\Theta = \Theta_1 + \Theta_2 + \Theta_3$ *where*

$$\begin{aligned}\Theta_1 = \quad & \langle \widehat{e}_w - e_w, (G_T - v_1)n\rangle \ + \langle \widehat{e}_u - e_u, (G_N - v_2)n\rangle \ - \langle \widehat{e}_\theta - e_\theta, (G_M - v_3)n\rangle \\ & + \langle \widehat{e}_M - e_M, (G_\theta - v_4)n\rangle - \langle \widehat{e}_N - e_N, (G_u - v_5)n\rangle - \langle \widehat{e}_T - e_T, (G_w - v_6)n\rangle,\end{aligned}$$

$$\begin{aligned}\Theta_2 = \quad & (e'_w, G_T - v_1) + (e'_u, G_N - v_2) - (e'_\theta, G_M - v_3) \\ & + (e'_M, G_\theta - v_4) - (e'_N, G_u - v_5) - (e'_T, G_w - v_6),\end{aligned}$$

and

$$\begin{aligned}\Theta_3 = \quad & (e_\theta + \kappa e_u - d^2 e_T, G_T - v_1) - (\kappa e_w + d^2 e_N, G_N - v_2) \\ & + (e_M, G_M - v_3) - (e_T, G_\theta - v_4) + (\kappa e_T, G_u - v_5) - (\kappa e_N, G_w - v_6)\end{aligned}$$

for all $(v_1, \ldots, v_6) \in \boldsymbol{V}_h^k$.

Proof. Adding and subtracting the term

$$\langle \widehat{e}_w, v_1 n\rangle + \langle \widehat{e}_u, v_2 n\rangle - \langle \widehat{e}_\theta, v_3 n\rangle + \langle \widehat{e}_M, v_4 n\rangle - \langle \widehat{e}_N, v_5 n\rangle - \langle \widehat{e}_T, v_6 n\rangle$$

to the original expression for Θ, we see that

$$\begin{aligned}\Theta = \quad & \langle \widehat{e}_w, (G_T - v_1)n\rangle \ + \langle \widehat{e}_u, (G_N - v_2)n\rangle - \langle \widehat{e}_\theta, (G_M - v_3)n\rangle \\ & + \langle \widehat{e}_M, (G_\theta - v_4)n\rangle - \langle \widehat{e}_N, (G_u - v_5)n\rangle - \langle \widehat{e}_T, (G_w - v_6)n\rangle \\ & + \langle \widehat{e}_w, v_1 n\rangle \qquad + \langle \widehat{e}_u, v_2 n\rangle \qquad - \langle \widehat{e}_\theta, v_3 n\rangle \\ & + \langle \widehat{e}_M, v_4 n\rangle \qquad - \langle \widehat{e}_N, v_5 n\rangle \qquad - \langle \widehat{e}_T, v_6 n\rangle.\end{aligned}$$

Rewriting the last six terms above by using the error equations (4.8a)-(4.8f), we obtain

$$\begin{aligned}\Theta = \quad & \langle \widehat{e}_w, (G_T - v_1)n\rangle + \langle \widehat{e}_u, (G_N - v_2)n\rangle - \langle \widehat{e}_\theta, (G_M - v_3)n\rangle \\ & + \langle \widehat{e}_M, (G_\theta - v_4)n\rangle - \langle \widehat{e}_N, (G_u - v_5)n\rangle - \langle \widehat{e}_T, (G_w - v_6)n\rangle \\ & + (e_w, v_1') + (e_u, v_2') - (e_\theta, v_3') + (e_M, v_4') - (e_N, v_5') - (e_T, v_6') \\ & - (e_\theta + \kappa e_u - d^2 e_T, v_1) + (\kappa e_w + d^2 e_N, v_2) \\ & - (e_M, v_3) + (e_T, v_4) - (\kappa e_T, v_5) + (\kappa e_N, v_6).\end{aligned}$$

Note that, by the definition of the Green's functions, we have

$$\begin{aligned} &(e_w, G_T') = -(e_w, \kappa G_N), && (e_u, G_N') = (e_u, \kappa G_T), \\ &(e_\theta, G_M') = -(e_\theta, G_T), && (e_M, G_\theta') = (e_M, G_M), \\ &(e_N, G_u') = (e_N, d^2 G_N - \kappa G_w), && (e_T, G_w') = (e_T, d^2 G_T + G_\theta - \kappa G_u). \end{aligned}$$

Inserting these equations into the last expression for Θ, and rearranging terms, we obtain

$$\begin{aligned}\Theta = \Theta_3 & + \langle \widehat{e}_w, (G_T - v_1)n\rangle + \langle \widehat{e}_u, (G_N - v_2)n\rangle - \langle \widehat{e}_\theta, (G_M - v_3)n\rangle \\ & + \langle \widehat{e}_M, (G_\theta - v_4)n\rangle - \langle \widehat{e}_N, (G_u - v_5)n\rangle - \langle \widehat{e}_T, (G_w - v_6)n\rangle \\ & - (e_w, (G_T - v_1)') \ - (e_u, (G_N - v_2)') \ + (e_\theta, (G_M - v_3)') \\ & - (e_M, (G_\theta - v_4)') \ + (e_N, (G_u - v_5)') \ + (e_T, (G_w - v_6)').\end{aligned}$$

It remains to show that

$$\begin{aligned}\Theta_1 + \Theta_2 = \quad & \langle \widehat{e}_w, (G_T - v_1)n\rangle + \langle \widehat{e}_u, (G_N - v_2)n\rangle - \langle \widehat{e}_\theta, (G_M - v_3)n\rangle \\ & + \langle \widehat{e}_M, (G_\theta - v_4)n\rangle - \langle \widehat{e}_N, (G_u - v_5)n\rangle - \langle \widehat{e}_T, (G_w - v_6)n\rangle \\ & - (e_w, (G_T - v_1)') \ - (e_u, (G_N - v_2)') \ + (e_\theta, (G_M - v_3)') \\ & - (e_M, (G_\theta - v_4)') \ + (e_N, (G_u - v_5)') \ + (e_T, (G_w - v_6)').\end{aligned}$$

This follows by integrating by parts on each of the last six terms. This completes the proof.

□

We are now ready to prove our representation result.

Proof. (Lemma 4.12) We begin by noting that, by the definition of the Green's functions, (4.13) and (4.14), we have

$$\Theta = \widehat{e}_z(x_i).$$

On the other hand, setting $\boldsymbol{v} = \boldsymbol{\Pi G}$ in Lemma (4.13), we obtain

$$\widehat{e}_z(x_i) = \Theta_1 + \Theta_2 + \Theta_3 \tag{4.33}$$

with

$$\begin{aligned}\Theta_1 = \quad & \langle \widehat{e}_w - e_w, \delta_{G_T} n\rangle \ + \langle \widehat{e}_u - e_u, \delta_{G_N} n\rangle \ - \langle \widehat{e}_\theta - e_\theta, \delta_{G_M} n\rangle \\ & + \langle \widehat{e}_M - e_M, \delta_{G_\theta} n\rangle - \langle \widehat{e}_N - e_N, \delta_{G_u} n\rangle - \langle \widehat{e}_T - e_T, \delta_{G_w} n\rangle,\end{aligned}$$

$$\begin{aligned}\Theta_2 = \quad & (e'_w, \delta_{G_T}) \ + (e'_u, \delta_{G_N}) - (e'_\theta, \delta_{G_M}) \\ & + (e'_M, \delta_{G_\theta}) - (e'_N, \delta_{G_u}) - (e'_T, \delta_{G_w}),\end{aligned}$$

and

$$\begin{aligned}\Theta_3 = \quad & (e_\theta + \kappa e_u - d^2 e_T, \delta_{G_T}) - (\kappa e_w + d^2 e_N, \delta_{G_N}) \\ & + (e_M, \delta_{G_M}) - (e_T, \delta_{G_\theta}) + (\kappa e_T, \delta_{G_u}) - (\kappa e_N, \delta_{G_w}).\end{aligned}$$

Clearly,

$$\Theta_3 = \Gamma_2. \tag{4.34}$$

By Lemma 4.10 with $\boldsymbol{v} = \boldsymbol{G}$ we have that

$$\Theta_1 = 0. \tag{4.35}$$

By the orthogonality property, (4.7a) and (4.7b), of the projection we have

$$\begin{aligned}\Theta_2 = \quad & (e'_w - (e'_w)_{k-1}, \delta_{G_T}) \ + (e'_u - (e'_u)_{k-1}, \delta_{G_N}) \ - (e'_\theta - (e'_\theta)_{k-1}, \delta_{G_M}) \\ & + (e'_M - (e'_M)_{k-1}, \delta_{G_\theta}) - (e'_N - (e'_N)_{k-1}, \delta_{G_u}) - (e'_T - (e'_T)_{k-1}, \delta_{G_w}).\end{aligned}$$

Since

$$\begin{aligned}e'_z - (e'_z)_{k-1} &= (z' - z'_h) - (z' - z'_h)_{k-1} \\ &= z' - (z')_{k-1} + (z'_h)_{k-1} - z'_h \\ &= z' - (z')_{k-1},\end{aligned}$$

we see that

$$\Theta_2 = \Gamma_1. \tag{4.36}$$

The result now follows from (4.33), (4.34), (4.35), and (4.36). □

Step 2: Proof of Theorem 4.6. We are now ready to prove Theorem 4.6.

By Lemma 4.12 we have that

$$|\widehat{e}_z(x_i)| \leq |\Gamma_1| + |\Gamma_2|. \tag{4.37}$$

The result then follows if we estimate each one of the terms appearing in Γ_1 and Γ_2. We will estimate one term from each expression since the remaining terms can be estimated similarly. From Γ_1, we estimate the term $(w' - (w')_{k-1}, \delta_{G_T})$. By the approximation properties of the

L^2-projection we get

$$\begin{aligned}(w' - (w')_{k-1}, \delta_{G_T}) &\leq \|w' - (w')_{k-1}\| \, \|\delta_{G_T}\| \\ &\leq C_{k-1} \, h^k |w'|_k \, \|\delta_{G_T}\| \\ &\leq C_{k-1} \, h^k |\mathbf{z}|_{k+1} \, \|\boldsymbol{\delta}_i^z\| \, .\end{aligned}$$

Estimating the remaining terms appearing in Γ_1 in a similar fashion we obtain

$$|\Gamma_1| \leq C_{k-1} \, h^k |\mathbf{z}|_{k+1} \, \|\boldsymbol{\delta}_i^z\| \, .$$

Finally, we show how to estimate the term $(e_\theta + \kappa e_u - d^2 e_T, \delta_{G_T})$ in Γ_2 since estimating the remaining terms is similar. Thus,

$$\begin{aligned}(e_\theta + \kappa e_u - d^2 e_T, \delta_{G_T}) &\leq (\|e_\theta\| + \|\kappa\|_\infty \|e_u\| + d^2 \|e_T\|) \, \|\delta_{G_T}\| \\ &\leq C \, \|\boldsymbol{e}\| \, \|\boldsymbol{\delta}_i^z\|\end{aligned}$$

since κ is bounded and $0 < d < 1$. This implies that

$$|\Gamma_2| \leq C \, \|\boldsymbol{e}\| \, \|\boldsymbol{\delta}_i^z\| \, .$$

Inserting the estimates of $|\Gamma_1|$ and $|\Gamma_2|$ into (4.37) completes the proof of Theorem 4.6.

4.7 Numerical results

In this section, we display numerical results to verify our theoretical findings. We solve the equations (1.3) and (1.4) in $\Omega = (0, 1)$ with $\kappa = 1$, together with the boundary conditions $w_D = u_D = \theta_N = 0$ on $\partial\Omega$. We take uniform loading in arc length, namely, $p = q = 1$ in Ω. Although, the theory has been carried out for variable curvature, we take a constant κ

so that we can compute the exact solution and produce history of convergence tables. This choice corresponds to a circular arch of thickness d.

The HDG method is defined by (4.1) whose numerical traces are given by (4.2) in which we take the stabilization function $\mathtt{S}$ to be constant on $\partial\Omega_h$.

We display our numerical results in Table 6 and Table 7. In Table 6, we present a history of convergence study for the stabilization function which is defined by setting

$$\alpha_\theta = \alpha_N = \alpha_T = \tau_1 = \tau_2 = \tau_3 = 1 \quad \text{on} \quad \partial\Omega_h.$$

In Table 7, we present analogous results with a different choice of the stabilization function, namely, we take

$$\alpha_\theta = \alpha_N = \alpha_T = 1, \quad \tau_1 = \tau_2 = \tau_3 = 0 \quad \text{on} \quad \partial\Omega_h.$$

In both tables, "mesh $= i$" means we employed a uniform mesh with 2^i elements to obtain the results of that particular row of the table. In Table 7, for the $k = 0$ column, "mesh $= i$" means we employed a uniform mesh with 2^{i+4} elements. We displayed results for $k = 0$ in this manner because it takes more refinements to reach the asymptotic regime, at least for this choice of the numerical traces. For polynomials degree $k = 0, 1, 2, 3$ we display the L^2-norm of the projection of the error, $\|\boldsymbol{\Pi e}\|$, the L^2-norm of the error, $\|\boldsymbol{e}\|$, and the error in the numerical traces, $\|\widehat{\boldsymbol{e}}\|_\infty$, defined by

$$\|\,\widehat{\boldsymbol{e}}\,\|_\infty := \max_{z\in\{T,N,M,\theta,u,w\}} \left(\max_{x\in\mathcal{E}_h} |(z - \widehat{z}_h)(x)| \right).$$

We also display numerical orders of convergence which are computed as follows. Let $\boldsymbol{e}(i)$ denote the error where a mesh with 2^i elements has been employed to obtain the HDG solution. As usual, the order of convergence, r_i, at level i is defined as $r_i := \log\big(\boldsymbol{e}(i-1)/\boldsymbol{e}(i)\big)/\log 2$.

Observe that the results displayed in both tables validate the superconvergence of order $k+2$ for $k \geq 1$, and optimal convergence for $k = 0$, of the projection of the error predicted by Theorem 4.4. We also see that the L^2-norm of the error converges optimally as was predicted by Theorem 4.5. The superconvergence of the numerical traces of order $2k + 1$ of Theorem 4.6 is also verified.

In these examples we took the thickness parameter $d = 10^{-2}$ but let us note that in the numerical experiments which we do not report here we observed similar results and exactly the same convergence orders when we took $d = 10^{-8}$. This verifies that the method is robust with respect to d and is free from locking as was predicted by our theoretical results in Sec. 4.5.

In Table 8, we compare the running time between DG and HDG methods for the same problem and we use 512 elements. We can see that the running time of HDG is much less than DG method.

4.8 Concluding remarks

We have shown that optimal HDG methods can be devised for Naghdi arches which are free from shear- and membrane-locking. We achieved this by a careful study of the relation between the definition of the numerical traces and the corresponding convergence properties of the methods. Key to our analysis was a new projection operator which is tailored to fit the structure of the numerical traces of the HDG method. We have shown that HDG solution superconverges to the projection of the exact solution for all the unknowns. This immediately results in optimal error estimates for all the unknowns. In this sense, the error analysis is

Table 6: $\alpha_\theta = \alpha_N = \alpha_T = \tau_1 = \tau_2 = \tau_3 = 1$.

mesh	$k=0$		$k=1$		$k=2$		$k=3$	
	$\|\boldsymbol{\Pi e}\|$	order	$\|\boldsymbol{\Pi e}\|$	order	$\|\boldsymbol{\Pi e}\|$	order	$\|\boldsymbol{\Pi e}\|$	order
3	2.38E-01	0.39	2.58E-03	2.87	5.07E-07	4.05	9.50E-10	5.18
4	1.56E-01	0.61	3.35E-04	2.94	3.14E-08	4.01	2.96E-11	5.01
5	9.06E-02	0.78	4.27E-05	2.97	1.96E-09	4.00	9.26E-13	5.00
6	4.88E-02	0.89	5.38E-06	2.99	1.23E-10	4.00	2.90E-14	5.00
	$\|\boldsymbol{e}\|$	order	$\|\boldsymbol{e}\|$	order	$\|\boldsymbol{e}\|$	order	$\|\boldsymbol{e}\|$	order
3	2.45E-01	0.44	1.12E-03	2.26	1.34E-05	3.08	5.04E-08	3.98
4	1.59E-01	0.63	2.15E-04	2.38	1.68E-06	3.00	3.17E-09	3.99
5	9.19E-02	0.79	4.62E-05	2.22	2.11E-07	2.99	1.98E-10	4.00
6	4.94E-02	0.90	1.09E-05	2.08	2.64E-08	3.00	1.24E-11	4.00
	$\|\widehat{\boldsymbol{e}}\|_\infty$	order	$\|\widehat{\boldsymbol{e}}\|_\infty$	order	$\|\widehat{\boldsymbol{e}}\|_\infty$	order	$\|\widehat{\boldsymbol{e}}\|_\infty$	order
3	2.38E-01	0.46	9.00E-04	2.38	1.56E-06	4.94	1.37E-10	7.01
4	1.53E-01	0.64	1.33E-04	2.76	4.95E-08	4.97	1.07E-12	7.01
5	8.80E-02	0.80	1.79E-05	2.89	1.56E-09	4.99	8.33E-15	7.00
6	4.72E-02	0.90	2.31E-06	2.95	4.90E-11	4.99	6.50E-17	7.00

Table 7: $\alpha_\theta = \alpha_N = \alpha_T = 1, \quad \tau_1 = \tau_2 = \tau_3 = 0.$

mesh	$k=0$		$k=1$		$k=2$		$k=3$	
	$\|\mathbf{\Pi} \boldsymbol{e}\|$	order	$\|\mathbf{\Pi} \boldsymbol{e}\|$	order	$\|\mathbf{\Pi} \boldsymbol{e}\|$	order	$\|\mathbf{\Pi} \boldsymbol{e}\|$	order
3	2.33E-01	0.66	3.70E-03	2.94	3.16E-07	4.19	1.19E-09	4.99
4	1.34E-01	0.79	4.66E-04	2.99	1.90E-08	4.06	3.72E-11	5.00
5	7.27E-02	0.88	5.83E-05	3.00	1.17E-09	4.02	1.16E-12	5.00
6	3.79E-02	0.94	7.30E-06	3.00	7.31E-11	4.00	3.64E-14	5.00
	$\|\boldsymbol{e}\|$	order	$\|\boldsymbol{e}\|$	order	$\|\boldsymbol{e}\|$	order	$\|\boldsymbol{e}\|$	order
3	2.33E-01	0.66	3.70E-03	2.94	3.16E-07	4.19	1.19E-09	4.99
4	1.34E-01	0.79	4.66E-04	2.99	1.90E-08	4.06	3.72E-11	5.00
5	7.27E-02	0.88	5.83E-05	3.00	1.17E-09	4.02	1.16E-12	5.00
6	3.80E-02	0.95	7.30E-06	3.00	7.31E-11	4.00	3.64E-14	5.00
	$\|\widehat{\boldsymbol{e}}\|_\infty$	order	$\|\widehat{\boldsymbol{e}}\|_\infty$	order	$\|\widehat{\boldsymbol{e}}\|_\infty$	order	$\|\widehat{\boldsymbol{e}}\|_\infty$	order
3	2.32E-01	0.66	3.69E-03	2.94	1.00E-07	4.95	1.22E-11	6.90
4	1.34E-01	0.79	4.65E-04	2.99	3.17E-09	4.98	9.83E-14	6.95
5	7.26E-02	0.88	5.83E-05	3.00	9.96E-11	4.99	7.81E-16	6.98
6	3.79E-02	0.94	7.29E-06	3.00	3.12E-12	5.00	6.15E-18	6.99

Table 8: Running Time between DG and HDG Methods

Polynomial Degree	$k = 1$	$k = 2$	$k = 3$
DG method	17.08s	54.81s	102.73s
HDG method	3.83s	7.75s	14.98s

simplified only to the study of the approximation properties of the projection operator.

This provides a powerful framework for devising locking-free HDG methods for more challenging problems arising in solid mechanics, such as the Naghdi shell model whose restriction from the 2-D model to 1-D results in the arch model we have considered here. This constitutes the subject of ongoing work.

5 Naghdi Type Shell Model

5.1 Notation

Let $\widetilde{\Omega} \subset \mathbb{R}^3$ be the middle surface of a shell of thickness 2ϵ. It is the image of a domain $\Omega \subset \mathbb{R}^2$ through a mapping Φ. In the following, we use an under-tilde to indicate the components of a 2-vector. Greek sub and super scripts take their values in $\{1, 2\}$. Summation rules with respect to repeated sub and super scripts will also be used. The coordinates $\underset{\sim}{x} = (x_1, x_2) \in \Omega$ then furnish the curvilinear coordinates on $\widetilde{\Omega}$. We assume that at any point on the surface, along the coordinate lines, the two tangential vectors $\boldsymbol{a}_\alpha = \partial\Phi/\partial x_\alpha$ are linearly independent. The unit vector $\boldsymbol{a}_3 = (\boldsymbol{a}_1 \times \boldsymbol{a}_2)/|\boldsymbol{a}_1 \times \boldsymbol{a}_2|$ is normal to $\widetilde{\Omega}$. The triple $\boldsymbol{a}_i$ furnishes the covariant basis on $\widetilde{\Omega}$. The contravariant basis $\boldsymbol{a}^i$ is defined by the relations $\boldsymbol{a}^\alpha \cdot \boldsymbol{a}_\beta = \delta^\alpha_\beta$ and $\boldsymbol{a}^3 = \boldsymbol{a}_3$, in which δ^α_β is the Kronecker delta. The metric tensor has the covariant components $a_{\alpha\beta} = \boldsymbol{a}_\alpha \cdot \boldsymbol{a}_\beta$. The determinant of this metric tensor is denoted by a. The contravariant components are given by $a^{\alpha\beta} = \boldsymbol{a}^\alpha \cdot \boldsymbol{a}^\beta$. The curvature tensor is defined by $b_{\alpha\beta} = \boldsymbol{a}_3 \cdot \partial_\beta \boldsymbol{a}_\alpha$. The mixed components are $b^\alpha_\beta = a^{\alpha\gamma} b_{\gamma\beta}$. The Christoffel symbols are defined by $\Gamma^\gamma_{\alpha\beta} = \boldsymbol{a}^\gamma \cdot \partial_\beta \boldsymbol{a}_\alpha$, which are symmetric with respect to the subscripts. The

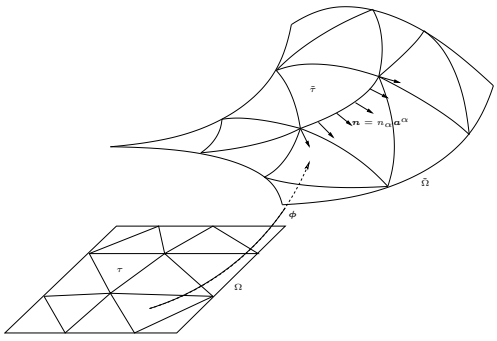

Figure 4: A triangularization of the shell surface.

covariant derivative of a vector or tensor is a higher order tensor. For example,

$$\sigma^{\alpha\beta}|_\gamma = \partial_\gamma \sigma^{\alpha\beta} + \Gamma^\alpha_{\gamma\lambda}\sigma^{\lambda\beta} + \Gamma^\beta_{\gamma\tau}\sigma^{\alpha\tau}, \quad \tau^\gamma_{\alpha|\beta} = \partial_\beta \tau^\gamma_\alpha + \Gamma^\gamma_{\lambda\beta}\tau^\lambda_\alpha - \Gamma^\tau_{\alpha\beta}\tau^\gamma_\tau,$$

$$u_{\alpha|\beta} = \partial_\beta u_\alpha - \Gamma^\gamma_{\alpha\beta}u_\gamma, \quad u^\alpha|_\beta = \partial_\beta u^\alpha + \Gamma^\alpha_{\gamma\beta}u^\gamma.$$

Product rules for differentiations, like $(\sigma^{\alpha\lambda}u_\lambda)|_\beta = \sigma^{\alpha\lambda}|_\beta u_\lambda + \sigma^{\alpha\lambda}u_{\lambda|\beta}$, are valid.

5.2 The Naghdi Type Shell

The Naghdi type shell model [86] determines the middle surface tangential displacement vector $u_\alpha \boldsymbol{a}^\alpha$, the transverse deflection vector $w\boldsymbol{a}_3$, and the normal fiber rotation vector $\theta_\alpha \boldsymbol{a}^\alpha$. A neater way to write the model is a 2D variational problem defined on a subspace H, determined by boundary conditions, of the multiple Sobolev space $\underset{\sim}{H}^1(\Omega) \times \underset{\sim}{H}^1(\Omega) \times H^1(\Omega)$. Here $\underset{\sim}{H}^1(\Omega) = [H^1(\Omega)]^2$. We let the tangential force density be $p^\alpha \boldsymbol{a}_\alpha$ and transverse force density be $p^3\boldsymbol{a}_3$. The model reads: Find $(\underset{\sim}{\theta}, \underset{\sim}{u}, w) \in H$, such that

$$\begin{aligned}\frac{1}{3}\int_\Omega a^{\alpha\beta\lambda\gamma}\rho_{\lambda\gamma}(\underset{\sim}{\theta}, \underset{\sim}{u}, w)\rho_{\alpha\beta}(\underset{\sim}{\phi}, \underset{\sim}{v}, z)\sqrt{a}d\underset{\sim}{x} + \epsilon^{-2}\int_\Omega a^{\alpha\beta\lambda\gamma}\gamma_{\lambda\gamma}(\underset{\sim}{u}, w)\gamma_{\alpha\beta}(\underset{\sim}{y}, z)\sqrt{a}d\underset{\sim}{x} \\ +\epsilon^{-2}\mu\int_\Omega a^{\alpha\beta}\tau_\beta(\underset{\sim}{\theta}, \underset{\sim}{u}, w)\tau_\alpha(\underset{\sim}{\phi}, \underset{\sim}{y}, z)\sqrt{a}d\underset{\sim}{x} = \int_\Omega (p^\alpha y_\alpha + p^3 z)\sqrt{a}d\underset{\sim}{x}\end{aligned} \tag{5.1}$$

for $\forall(\underset{\sim}{\phi}, \underset{\sim}{y}, z) \in H$. in which the fourth order two-dimensional contravariant tensor $a^{\alpha\beta\delta\gamma}$ is the elastic tensor of the shell, defined by

$$a^{\alpha\beta\delta\gamma} = \mu(a^{\alpha\delta}a^{\beta\gamma} + a^{\beta\delta}a^{\alpha\gamma}) + \lambda^* a^{\alpha\beta}a^{\delta\gamma}, \quad \text{with } \lambda^* = \frac{2\mu\lambda}{2\mu+\lambda}.$$

Here, λ and μ are the Lamé constants of the elastic material. This fourth order tensor is often given as $2\mu a^{\alpha\delta}a^{\beta\gamma} + \lambda^* a^{\alpha\beta}a^{\delta\gamma}$. But such a form loses certain symmetry. It is noted that when acting on a symmetric strain tensor the effect of this form is the same as that of the more formal one.

The compliance tensor of the shell defines the inverse operator of the elastic tensor, given by

$$a_{\alpha\beta\delta\gamma} = \frac{1}{2\mu}\left[\frac{1}{2}(a_{\alpha\gamma}a_{\beta\delta} + a_{\beta\gamma}a_{\alpha\delta}) - \frac{\lambda}{2\mu+3\lambda}a_{\alpha\beta}a_{\gamma\delta}\right]$$

We have

$$\sigma^{\alpha\beta} = a^{\alpha\beta\delta\gamma}\gamma_{\delta\gamma} \quad \Longleftrightarrow \quad \gamma_{\alpha\beta} = a_{\alpha\beta\delta\gamma}\sigma^{\delta\gamma}, \quad \text{if both } \sigma \text{ and } \gamma \text{ are symmetric.}$$

If, say, $\gamma_{12} \neq \gamma_{21}$, the left one could hold in which σ must be symmetric, while the right one must be broken. Note that $a^{\alpha\beta}|_\gamma = a_{\alpha\beta}|_\gamma = 0$. If the shell material has constant Lamé coefficients, we have $a^{\alpha\beta\gamma\delta}|_\tau = a_{\alpha\beta\gamma\delta}|_\tau = 0$.

For $(\underset{\sim}{\theta}, \underset{\sim}{u}, w) \in H$,

$$\begin{aligned}
\gamma_{\alpha\beta}(\underset{\sim}{u}, w) &= \frac{1}{2}(u_{\alpha|\beta} + u_{\beta|\alpha}) - b_{\alpha\beta}w, \\
\rho_{\alpha\beta}(\underset{\sim}{\theta}, \underset{\sim}{u}, w) &= \frac{1}{2}(\theta_{\alpha|\beta} + \theta_{\beta|\alpha}) - \frac{1}{2}(b^\lambda_\alpha u_{\lambda|\beta} + b^\gamma_\beta u_{\gamma|\alpha}) + c_{\alpha\beta}w, \\
\tau_\beta(\underset{\sim}{\theta}, \underset{\sim}{u}, w) &= b^\lambda_\beta u_\lambda + \theta_\beta + \partial_\beta w
\end{aligned} \tag{5.2}$$

are the membrane strain, bending strain and transverse shear strain engendered by the tangential displacement $\underset{\sim}{u}$, transverse displacement w, and normal fiber rotation $\underset{\sim}{\theta}$.

The Koiter model does not allow transverse shear. It can be derived by eliminating the variable $\underset{\sim}{\theta}$ with the vanishing shear condition $\tau_\beta(\underset{\sim}{\theta}, \underset{\sim}{u}, w) = b^\lambda_\beta u_\lambda + \theta_\beta + \partial_\beta w = 0$. The model determines $(\underset{\sim}{u}, w)$ in a subspace, still denoted by H, of $\underset{\sim}{H}{}^1(\Omega) \times H^2(\Omega)$, such that

$$\begin{aligned}
\frac{1}{3}\int_\Omega a^{\alpha\beta\lambda\gamma}\rho_{\lambda\gamma}(\underset{\sim}{u}, w)\rho_{\alpha\beta}(\underset{\sim}{y}, z)\sqrt{a}d\underset{\sim}{x} + \epsilon^{-2}\int_\Omega a^{\alpha\beta\lambda\gamma}\gamma_{\lambda\gamma}(\underset{\sim}{u}, w)\gamma_{\alpha\beta}(\underset{\sim}{y}, z)\sqrt{a}d\underset{\sim}{x} \\
= \int_\Omega (p^\alpha y_\alpha + p^3 z)\sqrt{a}d\underset{\sim}{x} \ \ \forall\ (\underset{\sim}{y}, z) \in H,
\end{aligned} \tag{5.3}$$

in which the elasticity tensor and the membrane (change of metric tensor) tensor are the same as that in the Naghdi model. The bending (change of curvature) tensor is changed to

$$\rho_{\alpha\beta}(\underset{\sim}{u}, w) = \partial^2_{\alpha\beta} w - \Gamma^\gamma_{\alpha\beta}\partial_\gamma w + b^\gamma_{\alpha|\beta} u_\gamma + b^\gamma_\alpha u_{\gamma|\beta} + b^\gamma_\beta u_{\gamma|\alpha} - c_{\alpha\beta} w. \tag{5.4}$$

Both (5.1) and (5.3) are well posed in their suitable spaces. Their solutions could be very elusive when $\epsilon \to 0$.

5.3 Green's theorem on surfaces

One may need to repeatedly use integration by parts. It seems advantageous to operate the calculation on the shell middle surface $\widetilde{\Omega}$, rather than on the two-dimensional domain Ω. For this purpose, we need Green's theorem, or divergence theorem, on the surface $\widetilde{\Omega}$. This theorem is a special case of the Stokes theorem regarding vector fields defined on a surface. Let $\tau \subset \Omega$ be an area element, which is mapped to $\widetilde{\tau} \subset \widetilde{\Omega}$ by Φ. Let $\boldsymbol{n} = n_\alpha \boldsymbol{a}^\alpha$ be the unit outward normal in the surface $\widetilde{\Omega}$ to the boundary of $\widetilde{\tau}$, denoted by $\partial\widetilde{\tau} = \Phi(\partial\tau)$, then for $\underset{\sim}{u} \in \underset{\sim}{H}(\mathrm{div}, \tau)$, one has

$$\int_{\widetilde{\tau}} u^\alpha|_\alpha dS = \int_\tau u^\alpha|_\alpha \sqrt{a} d\underset{\sim}{x} = \int_\tau (\sqrt{a} u^\alpha)_{,\alpha} d\underset{\sim}{x} = \int_{\partial\tau} \sqrt{a} u^\alpha n^{\partial\tau}_\alpha ds = \int_{\partial\widetilde{\tau}} u^\alpha n_\alpha ds. \tag{5.5}$$

In the equation, the left-most integral is taken with respect to the area measurement on S, while the right most integral is with respect to the arc length. The equalty of the third integral with the fourth integral is the classical divergence theorem on 2D Euclidean space. We often simply write the Green's theorem on surface as

$$\int_{\widetilde{\tau}} u^\alpha|_\alpha = \int_{\partial\widetilde{\tau}} u^\alpha n_\alpha.$$

This Green's theorem on surface can be proved by using the divergence theorem in the 3D space. Let $\widetilde{\tau}^\epsilon$ be a thin shell of thickness 2ϵ and mid-surface $\widetilde{\tau}$. We extend the 2D vector

field $\underset{\sim}{u}$ from $\widetilde{\tau}$ to a 3D vector field on the shell such that $u^\alpha(\underset{\sim}{x}, x_3) = u^\alpha(\underset{\sim}{x})$ and $u^3 = 0$. Then the 3D divergence theorem says that

$$\int_{\widetilde{\tau}^\epsilon} u^i\|_i = \int_{\partial\widetilde{\tau}^{\epsilon\pm}} u^i n_i^\pm + \int_{\partial\widetilde{\tau}^\epsilon} u^i n_i.$$

Here the last integral is taken on the shell lateral face where $n_3 = 0$. The first integral in the right hand side is on the upper and lower faces, and it is zero since $n_\alpha = 0$ there. The divergence in the left hand side is $u^i\|_i = u^\alpha_{,\alpha} + \Gamma^{*\gamma}_{\beta\gamma} u^\beta$, which is $u^\alpha|_\alpha$ when restricted on $\widetilde{\tau}$. The Green's theorem follows when we take the limit $\epsilon \to 0$.

This Green's theorem can also be proved by using the Stokes theorem on a surface, which says that for a vector field $\boldsymbol{v}$ on a surface $\widetilde{\tau}$ one has

$$\int_{\widetilde{\tau}} (\operatorname{curl} \boldsymbol{v}) \cdot \boldsymbol{n} = \int_{\partial\widetilde{\tau}} \boldsymbol{v} \cdot \boldsymbol{s}.$$

Here $\boldsymbol{n} = \boldsymbol{a}_3$ is the upward unit normal vector to the surface and $\boldsymbol{s}$ is the counterclockwise unit tangent vector to its boundary curve. Note that $\operatorname{curl} \boldsymbol{v} = \epsilon^{ijk} v_{j\|i} \boldsymbol{g}_k$. Thus $(\operatorname{curl} \boldsymbol{v}) \cdot \boldsymbol{n} = \epsilon^{\alpha\beta} v_{\beta|\alpha}$. From the Stokes theorem, we get

$$\int_{\widetilde{\tau}} \epsilon^{\alpha\beta} v_{\beta|\alpha} = \int_{\partial\widetilde{\tau}} v_\alpha s^\alpha.$$

This equation itself maybe called rotation theorem on surface. It is as important as the Green's (divergence) theorem. We then use the facts that $\epsilon^{\alpha\beta}|_\gamma = 0$ and $s^\alpha = \epsilon^{\beta\alpha} n_\beta$ to get

$$\int_{\widetilde{\tau}} [\epsilon^{\alpha\beta} v_\beta]|_\alpha = \int_{\partial\widetilde{\tau}} v_\alpha \epsilon^{\beta\alpha} n_\beta.$$

The Green's theorem on surface then follows by taking $u^\alpha = \epsilon^{\alpha\beta} v_\beta$. It is noted that the Stokes theorem can actually be proved by using the divergence theorem, or more accurately rotation theorem, on flat plane. A short cut of this observation and the second approach is

the third method to prove the Green's theorem on surface: dividing the surface into small elements, replacing each element by a flat planar segment, and using the Green's theorem on 2D flat plane, and doing the standard calculus process of refining approximations.

5.4 Naghdi model as a system of first order PDE's

For a clamped shell, the Naghdi model [86] determines $(\underset{\sim}{\theta}, \underset{\sim}{u}, w) \in H = \underset{\sim}{H}_0^1 \times \underset{\sim}{H}_0^1 \times H_0^1$ such that

$$\frac{1}{3}\int_{\widetilde{\Omega}} a^{\alpha\beta\lambda\gamma}\rho_{\lambda\gamma}(\underset{\sim}{\theta}, \underset{\sim}{u}, w)\rho_{\alpha\beta}(\underset{\sim}{\phi}, \underset{\sim}{v}, z) + \epsilon^{-2}\int_{\widetilde{\Omega}} a^{\alpha\beta\lambda\gamma}\gamma_{\lambda\gamma}(\underset{\sim}{u}, w)\gamma_{\alpha\beta}(\underset{\sim}{y}, z) + \epsilon^{-2}\mu\int_{\widetilde{\Omega}} a^{\alpha\beta}\tau_{\beta}(\underset{\sim}{\theta}, \underset{\sim}{u}, w)\tau_{\alpha}(\underset{\sim}{\phi}, \underset{\sim}{y}, z) = \int_{\widetilde{\Omega}} (p^{\alpha}y_{\alpha} + p^{3}z) \quad \forall\ (\underset{\sim}{\phi}, \underset{\sim}{y}, z) \in H.$$

There are many ways to write this variational equation in the strong partial differential equation form. One can also introduce the first derivatives as new variables and write the PDE's as of first order. It seems that for the Naghdi model a more natural way is introducing the scaled membrane stress tensor $\mathcal{M}$, scaled shear stress vector $\mathcal{S}$, and the bending tensor $\mathcal{B}$ by

$$\mathcal{M}^{\alpha\beta} = \epsilon^{-2}a^{\alpha\beta\lambda\gamma}\gamma_{\lambda\gamma}(\underset{\sim}{u}, w), \quad \mathcal{S}^{\alpha} = \epsilon^{-2}\mu a^{\alpha\beta}\tau_{\beta}(\underset{\sim}{\theta}, \underset{\sim}{u}, w), \quad \mathcal{B}^{\alpha\beta} = a^{\alpha\beta\lambda\gamma}\rho_{\lambda\gamma}(\underset{\sim}{\theta}, \underset{\sim}{u}, w).$$

The Naghdi model (5.1) can be then written as the following system of differential equa-

tions.

$$
\begin{aligned}
-\frac{1}{3}[\mathcal{B}^{\alpha\beta}]|_\beta + \mathcal{S}^\alpha &= 0,\\
\frac{1}{3}[\mathcal{B}^{\lambda\gamma} b^\alpha_\lambda]|_\gamma - \mathcal{M}^{\alpha\beta}|_\beta + b^\alpha_\lambda \mathcal{S}^\lambda &= p^\alpha,\\
c_{\alpha\beta}\frac{1}{3}\mathcal{B}^{\alpha\beta} - b_{\alpha\beta}\mathcal{M}^{\alpha\beta} - \mathcal{S}^\alpha|_\alpha &= p^3,\\
\gamma_{\alpha\beta}(\underset{\sim}{u}, w) - \epsilon^2 a_{\alpha\beta\lambda\gamma}\mathcal{M}^{\lambda\gamma} &= 0,\\
\mu\tau_\alpha(\underset{\sim}{\theta}, \underset{\sim}{u}, w) - \epsilon^2 a_{\alpha\beta}\mathcal{S}^\beta &= 0,\\
\rho_{\lambda\gamma}(\underset{\sim}{\theta}, \underset{\sim}{u}, w) - a_{\alpha\beta\lambda\gamma}\mathcal{B}^{\alpha\beta} &= 0.
\end{aligned}
\tag{5.6}
$$

This is a system of 13 equations for 13 two-variable functions.

5.5 Weak form of the first order PDE system

Let $\tau \subset \Omega$ be an element, and $\widetilde{\tau} = \Phi(\tau) \subset \widetilde{\Omega}$ be the mapped curvilinear surface element. We multiply the equations in (5.6) by test functions and integrate the product on $\widetilde{\tau}$. We pair the notations as

$$
\theta \iff \phi, \quad u \iff v, \quad w \iff z, \quad \mathcal{B} \iff \mathcal{C}, \quad \mathcal{M} \iff \mathcal{N}, \quad \mathcal{S} \iff \mathcal{T}.
$$

We multiply the equations, respectively, by ϕ_α, v_α, z, $\mathcal{N}^{\alpha\beta}$, $\mathcal{T}^\alpha$, and $\mathcal{C}^{\lambda\gamma}$. The first equation becomes

$$
-\frac{1}{3}\int_{\widetilde{\tau}} [\mathcal{B}^{\alpha\beta}]|_\beta \phi_\alpha + \int_{\widetilde{\tau}} \mathcal{S}^\alpha \phi_\alpha = 0.
$$

By the Green's theorem on surfaces, this can be written as

$$
\frac{1}{3}\int_{\widetilde{\tau}} \mathcal{B}^{\alpha\beta}\frac{\phi_{\alpha|\beta} + \phi_{\beta|\alpha}}{2} + \int_{\widetilde{\tau}} \mathcal{S}^\alpha \phi_\alpha - \frac{1}{3}\int_{\partial\widetilde{\tau}} \mathcal{B}^{\alpha\beta}\phi_\alpha n_\beta = 0. \tag{5.7}
$$

We time the second equation of (5.6) by v_α and integrate and apply Green's theorem. The second one can be written as

$$-\frac{1}{3}\int_{\widetilde{\tau}}\mathcal{B}^{\alpha\beta}\frac{1}{2}(b^\lambda_\alpha u_{\lambda|\beta}+b^\gamma_\beta u_{\gamma|\alpha})+\int_{\widetilde{\tau}}\mathcal{M}^{\alpha\beta}\frac{v_{\alpha|\beta}+v_{\beta|\alpha}}{2}+\int_{\widetilde{\tau}}\mathcal{S}^\alpha b^\lambda_\alpha v_\lambda \\ +\frac{1}{3}\int_{\partial\widetilde{\tau}}\mathcal{B}^{\alpha\beta}b^\gamma_\alpha v_\gamma n_\beta-\int_{\partial\widetilde{\tau}}\mathcal{M}^{\alpha\beta}v_\alpha n_\beta=\int_{\widetilde{\tau}}p^\alpha v_\alpha. \quad (5.8)$$

By timing z to the third equation, integrate, and integrate by parts, the third equation becomes

$$\frac{1}{3}\int_{\widetilde{\tau}}\mathcal{B}^{\alpha\beta}c_{\alpha\beta}z \quad - \quad \int_{\widetilde{\tau}}\mathcal{M}^{\alpha\beta}b_{\alpha\beta}z \quad + \quad \int_{\widetilde{\tau}}\mathcal{S}^\alpha\partial_\alpha z \quad - \quad \int_{\partial\widetilde{\tau}}\mathcal{S}^\alpha n_\alpha z \qquad = \qquad \int_{\widetilde{\tau}}p^3 z. \quad (5.9)$$

Summing up these equations, invoking the definition (5.2), we have

$$\frac{1}{3}\int_{\widetilde{\tau}}\mathcal{B}^{\alpha\beta}\rho_{\alpha\beta}(\underset{\sim}{\phi},\underset{\sim}{v},z)+\int_{\widetilde{\tau}}\mathcal{M}^{\alpha\beta}\gamma_{\alpha\beta}(\underset{\sim}{v},z)+\int_{\widetilde{\tau}}\mathcal{S}^\alpha\tau_\alpha(\underset{\sim}{\phi},\underset{\sim}{v},z) \\ -\frac{1}{3}\int_{\partial\widetilde{\tau}}\mathcal{B}^{\alpha\beta}\phi_\alpha n_\beta+\frac{1}{3}\int_{\partial\widetilde{\tau}}\mathcal{B}^{\alpha\beta}b^\gamma_\alpha v_\gamma n_\beta-\int_{\partial\widetilde{\tau}}\mathcal{M}^{\alpha\beta}v_\alpha n_\beta-\int_{\partial\widetilde{\tau}}\mathcal{S}^\alpha n_\alpha z \\ =\int_{\widetilde{\tau}}p^\alpha v_\alpha+\int_{\widetilde{\tau}}p^3 z. \quad (5.10)$$

The last three equations in (5.6) can be written as

$$\frac{1}{3}\int_{\widetilde{\tau}}\mathcal{C}^{\alpha\beta}\rho_{\alpha\beta}(\underset{\sim}{\theta},\underset{\sim}{u},w)+\int_{\widetilde{\tau}}\mathcal{N}^{\alpha\beta}\gamma_{\alpha\beta}(\underset{\sim}{u},w)+\mu\int_{\widetilde{\tau}}\mathcal{T}^\alpha\tau_\alpha(\underset{\sim}{\theta},\underset{\sim}{u},w) \\ -\frac{1}{3}\int_{\widetilde{\tau}}a_{\alpha\beta\lambda\gamma}\mathcal{C}^{\alpha\beta}\mathcal{B}^{\lambda\gamma}-\epsilon^2\int_{\widetilde{\tau}}a_{\alpha\beta\lambda\gamma}\mathcal{N}^{\alpha\beta}\mathcal{M}^{\lambda\gamma}-\epsilon^2\int_{\widetilde{\tau}}a_{\alpha\beta}\mathcal{T}^\beta\mathcal{S}^\alpha=0. \quad (5.11)$$

When one do this for each element on the shell, and add up, one would need to resolve the border terms represented by

$$\sum_{\tau\in\mathcal{T}_h}\left[-\frac{1}{3}\int_{\partial\widetilde{\tau}}\mathcal{B}^{\alpha\beta}\phi_\alpha n_\beta+\frac{1}{3}\int_{\partial\widetilde{\tau}}\mathcal{B}^{\alpha\beta}b^\gamma_\alpha v_\gamma n_\beta-\int_{\partial\widetilde{\tau}}\mathcal{M}^{\alpha\beta}v_\alpha n_\beta-\int_{\partial\widetilde{\tau}}\mathcal{S}^\alpha n_\alpha z\right]$$

On an interior border $\widetilde{e} \in \mathcal{E}_h^0$, using the fact that $n_\beta^+ + n_\beta^- = 0$, the first term can be written as

$$\int_{\widetilde{e}} [\mathcal{B}^{\alpha\beta}\phi_\alpha n_\beta]^+ + [\mathcal{B}^{\alpha\beta}\phi_\alpha n_\beta]^- = \int_{\widetilde{e}} [\![\mathcal{B}^{\alpha\beta}]\!]_{n_\beta} \{\!\{\phi_\alpha\}\!\} + \int_{\widetilde{e}} \{\!\{\mathcal{B}^{\alpha\beta}\}\!\} [\![\phi_\alpha]\!]_{n_\beta}.$$

The jump and average are defined as

$$[\![\mathcal{B}^{\alpha\beta}]\!]_{n_\beta} = [\mathcal{B}^{\alpha\beta} n_\beta]^+ + [\mathcal{B}^{\alpha\beta} n_\beta]^-, \quad \{\!\{\phi_\alpha\}\!\} = \frac{[\phi_\alpha]^+ + [\phi_\alpha]^-}{2}, \quad [\![\phi_\alpha]\!]_{n_\beta} = [\phi_\alpha n_\beta]^+ + [\phi_\alpha n_\beta]^-.$$

APPENDIX A: PROOF OF THEOREM 2.2

In this appendix we prove Theorem 4.1 which guarantees the existence and uniqueness of the DG approximation. We will use the following lemma.

Lemma A. *Let $k, \ell, s,$ and t be non-negative integers. Let $f \in \mathcal{P}^k([a,b])$ and $g \in \mathcal{P}^\ell([a,b])$ be such that*

$$f(a) = g(a) = 0. \tag{A.1}$$

Suppose that

$$s, t \geq \max\{k, \ell\}, \tag{A.2}$$

and that

$$\begin{aligned} \mathsf{P}_s(g' + \alpha f) &= 0, \\ \mathsf{P}_t(f' - \alpha g) &= 0, \end{aligned} \tag{A.3}$$

where α is a function in $L^\infty([a,b])$ and $\mathsf{P}_\star$ denotes the L^2-orthogonal projection into $\mathcal{P}^\star([a,b])$. Then $f = g = 0$ in $[a,b]$ if

(a) α is identically equal to a constant, or

(b) α is not identically equal to a constant and

$$b - a \leq \frac{1}{2\left\|\alpha - \overline{\alpha}\right\|_{L^\infty([a,b])}} \tag{A.4}$$

where $\overline{\alpha}$ denotes the average value of α over the interval $[a,b]$.

Proof. Suppose that $s \geq t$, then by (A.3), $\mathsf{P}_t(g' + \alpha f) = 0$ and $\mathsf{P}_t(f' - \alpha g) = 0$. Since $t \geq \max\{k, \ell\}$ we see that

$$g' + \mathsf{P}_t(\alpha f) = 0, \tag{A.5a}$$

$$f' - \mathsf{P}_t(\alpha g) = 0, \tag{A.5b}$$

pointwise on $[a,b]$. Multiplying (A.5a) by g and (A.5b) by f we get

$$\frac{1}{2}(g^2)' + g\mathsf{P}_t(\alpha f) = 0, \qquad \frac{1}{2}(f^2)' - f\mathsf{P}_t(\alpha g) = 0,$$

and hence

$$\frac{1}{2}(g^2+f^2)' = f\mathsf{P}_t(\alpha g) - g\mathsf{P}_t(\alpha f) = f\mathsf{P}_t((\alpha-\overline{\alpha})g) - g\mathsf{P}_t((\alpha-\overline{\alpha})f) \tag{A.6}$$

since $-f\mathsf{P}_t(\overline{\alpha}g)+g\mathsf{P}_t(\overline{\alpha}f) = 0$ by (A.2). Integrating both sides of (A.6) from a to an arbitrary x in $[a,b]$, and using (A.1), we obtain

$$\frac{1}{2}(g^2+f^2)(x) = T_1(x) + T_2(x)$$

where

$$T_1(x) = \int_a^x f(s)\mathsf{P}_t((\alpha-\overline{\alpha})g)(s)\,ds, \qquad T_2(x) = -\int_a^x g(s)\mathsf{P}_t((\alpha-\overline{\alpha})f)(s)\,ds.$$

By Cauchy-Schwarz inequality

$$\begin{aligned}|T_1(x)| &\le \|f\|_{L^2([a,b])}\,\|(\alpha-\overline{\alpha})g\|_{L^2([a,b])} \\ &\le \|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\,\|f\|_{L^2([a,b])}\,\|g\|_{L^2([a,b])}\,.\end{aligned}$$

Similarly,

$$|T_2(x)| \le \|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\,\|f\|_{L^2([a,b])}\,\|g\|_{L^2([a,b])}\,,$$

and hence

$$\frac{1}{2}(g^2+f^2)(x) \le 2\,\|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\,\|f\|_{L^2([a,b])}\,\|g\|_{L^2([a,b])}\,.$$

Integrating both sides over $x\in[a,b]$ implies

$$\begin{aligned}\frac{1}{2}(\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])}) &\le 2(b-a)\,\|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\,\|f\|_{L^2([a,b])}\,\|g\|_{L^2([a,b])} \\ &\le (b-a)\,\|\alpha-\overline{\alpha}\|_{L^\infty([a,b])}\,(\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])})\end{aligned}$$

by Young's inequality. Thus,

$$\left[\frac{1}{2} - (b-a)\left\|\alpha - \overline{\alpha}\right\|_{L^\infty([a,b])}\right]\left(\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])}\right) \leq 0. \tag{A.7}$$

Now, if α is identically constant on $[a,b]$ then $\overline{\alpha} = \alpha$ and the result follows since in such a case (A.7) implies $\|f\|^2_{L^2([a,b])} + \|g\|^2_{L^2([a,b])} = 0$. If α is not identically constant on $[a,b]$ then we reach the same conclusion by (A.4).

The same conclusion can be reached if $s \leq t$ by following a similar argument. This completes the proof. □

We are now ready to prove Theorem 2.2.

Theorem *2.2.* Due to the linearity of the problem it suffices to show that the only solution to (2.2) with

$$p = q = 0 \quad \text{on } \Omega,$$

and

$$w_0 = w_1 = u_0 = u_1 = \theta_0 = \theta_1 = 0,$$

is

$$w_h = u_h = \theta_h = M_h = N_h = T_h = 0 \quad \text{on } \Omega_h.$$

In this case, (4.15) takes the form $\Theta_{interior} + \Theta_{jumps} = 0$, where

$$\Theta_{interior} = d^2(T_h, T_h)_{\Omega_h} + d^2(N_h, N_h)_{\Omega_h} + (M_h, M_h)_{\Omega_h},$$

and

$$\Theta_{jumps} = -\sum_{e\in\mathcal{E}_h}\Big(C_{16}[\![T_h]\!]^2 + C_{25}[\![N_h]\!]^2 + C_{34}[\![M_h]\!]^2 + C_{43}[\![\theta_h]\!]^2 + C_{52}[\![u_h]\!]^2 + C_{61}[\![w_h]\!]^2\Big)(e).$$

By assumption (2.8), this implies $T_h = N_h = M_h = 0 \quad$ on Ω_h. We also have that $[\![\theta_h]\!] = [\![u_h]\!] = [\![w_h]\!] = 0 \quad$ on $\mathcal{E}_h$, and hence θ_h, u_h, and w_h are continuous functions over Ω. Consequently, by (4.2), (2.4), and (2.5), $\widehat{\theta}_h = \theta_h$, $\widehat{u}_h = u_h$, $\widehat{w}_h = w_h$, on $\mathcal{E}_h$. Equation (2.2c) can then be written as $-(\theta_h, v_3')_{\Omega_h} + \langle \theta_h, v_3 \rangle_{\mathcal{E}_h} = 0$. Upon integration by parts we get $(\theta_h', v_3)_{\Omega_h} = 0$ for all $v_3 \in V_h^{k_3}$. Since $\theta_h \in V_h^{k_4}$ and $k_3 \geq k_4 - 1$ by assumption (2.10), we see that $\theta_h \equiv 0$ on Ω_h.

The remaining DG equations can now be written as

$$
\begin{aligned}
-(w_h, v_1')_{\Omega_h} + \langle w_h, [\![v_1]\!] \rangle_{\mathcal{E}_h} + (\kappa u_h, v_1)_{\Omega_h} &= 0, \\
-(u_h, v_2')_{\Omega_h} + \langle u_h, [\![v_2]\!] \rangle_{\mathcal{E}_h} - (\kappa w_h, v_2)_{\Omega_h} &= 0,
\end{aligned}
$$

for all $(v_1, v_2) \in V_h^{k_1} \times V_h^{k_2}$. Upon integrating by parts these equations become $(w_h' + \kappa u_h, v_1)_{\Omega_h} = 0$, $(u_h' - \kappa w_h, v_2)_{\Omega_h} = 0$, and hence $\mathsf{P}_{k_1}(w_h' + \kappa u_h) = 0$, and $\mathsf{P}_{k_2}(u_h' - \kappa w_h) = 0$ in Ω_h. If we apply Lemma A with

$$
g = w_h,\ f = u_h,\ \alpha = \kappa,\ k = k_5,\ \ell = k_6,\ s = k_1,\ t = k_2,\ a = x_0,\ b = x_1,
$$

we see that $w_h = u_h = 0$ on I_1 by (2.10) and (2.11), since $w_h(0) = w_0 = 0$ and $u_h(0) = u_0 = 0$. In particular, we get that $w_h(x_1) = u_h(x_1) = 0$, and hence we can apply Lemma A once more with $a = x_1$, $b = x_2$ and deduce that $w_h = u_h = 0$ on I_2. Similarly, we can prove that $w_h = u_h = 0$ on Ω_h. This completes the proof. $\square$

APPENDIX B:
PROOF OF CHARACTERIZATION THEOREM

The conservativity conditions for the HDG methods for Naghdi arches are

$$\langle \widehat{\theta}_h, \mathsf{m}\, n\rangle = \langle \theta_N, \mathsf{m}\, n\rangle_{\partial\Omega}, \tag{B.1a}$$

$$\langle \widehat{N}_h, \mathsf{u}\, n\rangle = 0, \tag{B.1b}$$

$$\langle \widehat{T}_h, \mathsf{w}\, n\rangle = 0, \tag{B.1c}$$

hold for all

$$(\mathsf{m}, \mathsf{u}, \mathsf{w}) \in L^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h) \times L_0^2(\mathcal{E}_h).$$

The lagrange multiplies are approximations at the nodes to w, u, and M, which are denote by w_h, u_h, and μ_h.

There are five local solvers, we label their equations as (w), (u), (μ), (p), (q), each of which contains six subequations and three more equations designing their numerical traces.

Since $T_h = \mathcal{T}\omega_h + \mathcal{T}w_D + \mathcal{T}u_h + \mathcal{T}u_D + \mathcal{T}\mu_h + \mathcal{T}p + \mathcal{T}q$ and similarly for N_h and θ_h, to

prove a characterization result we need to work out expressions for

$$\langle \widehat{\Theta}_w, \mathsf{m}\, n \rangle \tag{B.2a}$$

$$\langle \widehat{\Theta}_{w_D}, \mathsf{m}\, n \rangle \tag{B.2b}$$

$$\langle \widehat{\Theta}_u, \mathsf{m}\, n \rangle \tag{B.2c}$$

$$\langle \widehat{\Theta}_{u_D}, \mathsf{m}\, n \rangle \tag{B.2d}$$

$$\langle \widehat{\Theta}_\mu, \mathsf{m}\, n \rangle \tag{B.2e}$$

$$\langle \widehat{\Theta}_p, \mathsf{m}\, n \rangle \tag{B.2f}$$

$$\langle \widehat{\Theta}_q, \mathsf{m}\, n \rangle \tag{B.2g}$$

$$\langle \widehat{\mathcal{N}}_w, \mathsf{u}\, n \rangle \tag{B.3a}$$

$$\langle \widehat{\mathcal{N}}_{w_D}, \mathsf{u}\, n \rangle \tag{B.3b}$$

$$\langle \widehat{\mathcal{N}}_u, \mathsf{u}\, n \rangle \tag{B.3c}$$

$$\langle \widehat{\mathcal{N}}_{u_D}, \mathsf{u}\, n \rangle \tag{B.3d}$$

$$\langle \widehat{\mathcal{N}}_\mu, \mathsf{u}\, n \rangle \tag{B.3e}$$

$$\langle \widehat{\mathcal{N}}_p, \mathsf{u}\, n \rangle \tag{B.3f}$$

$$\langle \widehat{\mathcal{N}}_q, \mathsf{u}\, n \rangle \tag{B.3g}$$

$$\langle \widehat{\mathcal{T}}_w, \mathsf{w}\, n\rangle \tag{B.4a}$$

$$\langle \widehat{\mathcal{T}}_{w_D}, \mathsf{w}\, n\rangle \tag{B.4b}$$

$$\langle \widehat{\mathcal{T}}_u, \mathsf{w}\, n\rangle \tag{B.4c}$$

$$\langle \widehat{\mathcal{T}}_{u_D}, \mathsf{w}\, n\rangle \tag{B.4d}$$

$$\langle \widehat{\mathcal{T}}_\mu, \mathsf{w}\, n\rangle \tag{B.4e}$$

$$\langle \widehat{\mathcal{T}}_p, \mathsf{w}\, n\rangle \tag{B.4f}$$

$$\langle \widehat{\mathcal{T}}_q, \mathsf{w}\, n\rangle \tag{B.4g}$$

Proof of (B.2a), we begin by writing $\langle \widehat{\Theta}_w, \mathsf{m}\, n\rangle = \langle \widehat{\Theta}_w - \Theta_w, \mathsf{m}\, n\rangle + \langle \Theta_w, \mathsf{m}\, n\rangle$ taking $\mathsf{u} = \mathsf{m}$ and $v_4 = \Theta_w$ in the local solver $(\mu) \Longrightarrow$

$$\langle \Theta_w, \mathsf{m}\, n\rangle = (\mathcal{M}_\mathsf{m}, (\Theta_w)') + (\mathcal{T}_\mathsf{m}, \Theta_w)$$

IBP $\Longrightarrow$

$$(\mathcal{M}_\mathsf{m}, (\Theta_w)') = \langle 1, \mathcal{M}_\mathsf{m}(\Theta_w)n\rangle - ((\mathcal{M}_\mathsf{m})', \Theta_w)$$

Using the local solver (w) with $v_1 = \mathcal{T}_\mathsf{m} \Longrightarrow$

$$(\Theta_w, \mathcal{T}_\mathsf{m}) = (\mathcal{W}_w, (\mathcal{T}_\mathsf{m})') - \langle w, (\mathcal{T}_\mathsf{m})n\rangle - (\kappa\mathcal{U}_\omega, \mathcal{T}_\mathsf{m}) + d^2(\mathcal{T}_w, \mathcal{T}_\mathsf{m})$$

Then combine these three together$\Longrightarrow$

$$\langle \Theta_w, \mathsf{m}\, n\rangle = \langle 1, \mathcal{M}_\mathsf{m}(\Theta_w)n\rangle - ((\mathcal{M}_\mathsf{m})', \Theta_w) + (\mathcal{W}_w, (\mathcal{T}_\mathsf{m})') - \langle w, (\mathcal{T}_\mathsf{m})n\rangle - (\kappa\mathcal{U}_\omega, \mathcal{T}_\mathsf{m} + d^2(\mathcal{T}_w, \mathcal{T}_\mathsf{m})$$

IBP $\Longrightarrow$

$$(\mathcal{W}_w, (\mathcal{T}_\mathsf{m})') = \langle 1, \mathcal{W}_w(\mathcal{T}_\mathsf{m})n\rangle - (\mathcal{T}_\mathsf{m}, (\mathcal{W}_w)')$$

taking $v_3 = \mathcal{M}_{\mathsf{m}}$ in the local solver $(w) \Longrightarrow$

$$-(\Theta_w, (\mathcal{M}_{\mathsf{m}})') = -\langle 1, \widehat{\Theta}_w(\mathcal{M}_{\mathsf{m}})n\rangle + (\mathcal{M}_w, \mathcal{M}_{\mathsf{m}})$$

Then we have $\Longrightarrow$

$$\begin{aligned}
\langle \Theta_w, \mathsf{m}n\rangle =& \langle 1, \mathcal{M}_{\mathsf{m}}(\Theta_w)n\rangle && -\langle w, (\mathcal{T}_{\mathsf{m}})n\rangle \\
& -\langle 1, \mathcal{M}_{\mathsf{m}}(\widehat{\Theta}_{\mathsf{m}})n\rangle && +(\mathcal{M}_w, \mathcal{M}_{\mathsf{m}}) \\
& +\langle 1, \mathcal{W}_w(\mathcal{T}_{\mathsf{m}})n\rangle && -(\mathcal{T}_{\mathsf{m}}, (\mathcal{W}_w)') \\
& -(\kappa\mathcal{U}_\omega, \mathcal{T}_{\mathsf{m}}) && +d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}})
\end{aligned}$$

which can be written as

$$\begin{aligned}
\langle \Theta_w, \mathsf{m}n\rangle =& d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) && +(\mathcal{M}_w, \mathcal{M}_{\mathsf{m}}) \\
& -\langle \widehat{\Theta}_w - \Theta_w, (\mathcal{M}_{\mathsf{m}})n\rangle && -\langle w, (\mathcal{T}_{\mathsf{m}})n\rangle \\
& +\langle 1, \mathcal{W}_w(\mathcal{T}_{\mathsf{m}})n\rangle && \\
& -(\kappa\mathcal{U}_\omega, \mathcal{T}_{\mathsf{m}}) && -(\mathcal{T}_{\mathsf{m}}, (\mathcal{W}_w)')
\end{aligned}$$

taking $\mu = \mathsf{m}$ and $v_6 = \mathcal{W}_w$ in the local solver $(\mu) \Longrightarrow$

$$-(\mathcal{T}_{\mathsf{m}}, (\mathcal{W}_w)') = -\langle \widehat{\mathcal{T}}_{\mathsf{m}}, (\mathcal{W}_w)n\rangle - (\kappa\mathcal{N}_{\mathsf{m}}, \mathcal{W}_w)$$

Then we have $\Longrightarrow$

$$\begin{aligned}
\langle \Theta_w, \mathsf{m}n\rangle =& d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) && +(\mathcal{M}_w, \mathcal{M}_{\mathsf{m}}) \\
& -\langle \widehat{\Theta}_w - \Theta_w, (\mathcal{M}_{\mathsf{m}})n\rangle && -\langle w, (\mathcal{T}_{\mathsf{m}})n\rangle \\
& -\langle \widehat{\mathcal{T}}_{\mathsf{m}} - \mathcal{T}_{\mathsf{m}}, (\mathcal{W}_w)n\rangle && \\
& -(\kappa\mathcal{U}_\omega, \mathcal{T}_{\mathsf{m}}) && -(\kappa\mathcal{N}_{\mathsf{m}}, \mathcal{W}_w)
\end{aligned}$$

taking $\mu = \mathsf{m}$ and $v_5 = \mathcal{U}_w$ in the local solver $(\mu) \Longrightarrow$

$$\begin{aligned} -(\kappa\mathcal{T}_{\mathsf{m}}, \mathcal{U}_w) &= (\mathcal{N}_{\mathsf{m}}, (\mathcal{U}_w)') && +\langle\widehat{\mathcal{N}}_{\mathsf{m}}, \mathcal{U}_w n\rangle \\ &= \langle\mathcal{N}_{\mathsf{m}}, (\mathcal{U}_w n\rangle && -(\mathcal{U}_w, (\mathcal{N}_{\mathsf{m}})') - \langle\widehat{\mathcal{N}}_{\mathsf{m}}, \mathcal{U}_w n\rangle \\ &= -\langle\widehat{\mathcal{N}}_{\mathsf{m}} - \mathcal{N}_{\mathsf{m}}, \mathcal{U}_w n\rangle && -(\mathcal{U}_w, (\mathcal{N}_{\mathsf{m}})') \end{aligned}$$

taking $v_2 = \mathcal{N}_{\mathsf{m}}$ in the local solver $(w) \Longrightarrow$

$$-(\mathcal{U}_w, (\mathcal{N}_{\mathsf{m}})') = d^2(\mathcal{N}_w, \mathcal{N}_{\mathsf{m}}) + (\kappa\mathcal{W}_w, \mathcal{N}_{\mathsf{m}})$$

Thus $\Longrightarrow$

$$\begin{aligned} \langle\Theta_w, \mathsf{m}n\rangle = & d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) && +d^2(\mathcal{N}_w, \mathcal{N}_{\mathsf{m}}) + (\mathcal{M}_w, \mathcal{M}_{\mathsf{m}}) \\ & -\langle w, \mathcal{T}_{\mathsf{m}} n\rangle && +\langle\widehat{\Theta}_w - \Theta_w, (\mathsf{m} - \mathcal{M}_{\mathsf{m}})n\rangle \\ & && -\langle\widehat{\mathcal{N}}_{\mathsf{m}} - \mathcal{N}_{\mathsf{m}}, \mathcal{U}_w n\rangle \\ & && -\langle\widehat{\mathcal{T}}_{\mathsf{m}} - \mathcal{T}_{\mathsf{m}}, \mathcal{W}_w n\rangle \end{aligned}$$

Then we have $\Longrightarrow$

$$\begin{aligned} \langle\Theta_w, \mathsf{m}n\rangle = & d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) && +d^2(\mathcal{N}_w, \mathcal{N}_{\mathsf{m}}) + (\mathcal{M}_w, \mathcal{M}_{\mathsf{m}}) \\ & -\langle w, \mathcal{T}_{\mathsf{m}} n\rangle && -\langle\widehat{\Theta}_w - \Theta_w, \mathcal{M}_{\mathsf{m}} n\rangle \\ & && -\langle\widehat{\mathcal{N}}_{\mathsf{m}} - \mathcal{N}_{\mathsf{m}}, \mathcal{U}_w n\rangle \\ & && -\langle\widehat{\mathcal{T}}_{\mathsf{m}} - \mathcal{T}_{\mathsf{m}}, \mathcal{W}_w n\rangle \end{aligned}$$

To prove an identity for $\langle\Theta_{w_D}, \mathsf{m}n\rangle$ we need to further work on the energy terms.

taking $\mu = \mathsf{m}$ and $v_1 = \mathcal{T}_w$ in the local solver $(\mu) \Longrightarrow$

$$d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) = -(\mathcal{W}_{\mathsf{m}}, (\mathcal{T}_w)') + (\Theta_{\mathsf{m}}, \mathcal{T}_w) + (\kappa\mathcal{U}_{\mathsf{m}}, \mathcal{T}_w)$$

IBP $\Longrightarrow$

$$-(\mathcal{W}_{\mathsf{m}}, (\mathcal{T}_w)') = -\langle 1, (\mathcal{W}_{\mathsf{m}})(\mathcal{T}_w)n\rangle + (\mathcal{T}_w, (\mathcal{W}_{\mathsf{m}})')$$

taking $v_4 = \Theta_{\mathsf{m}}$ in the local solver $(w) \Longrightarrow$

$$(\Theta_{\mathsf{m}}, \mathcal{T}_w) = -(\mathcal{M}_w, (\Theta_{\mathsf{m}})')$$

Then $\Longrightarrow$

$$d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) = -\langle 1, (\mathcal{W}_{\mathsf{m}})(\mathcal{T}_w)n\rangle + (\mathcal{T}_w, (\mathcal{W}_{\mathsf{m}})') - (\mathcal{M}_w, (\Theta_{\mathsf{m}})') + (\kappa\mathcal{U}_{\mathsf{m}}, \mathcal{T}_w)$$

IBP $\Longrightarrow$

$$-(\mathcal{M}_w, (\Theta_{\mathsf{m}})') = -\langle 1, (\mathcal{M}_w)(\Theta_{\mathsf{m}})n\rangle + (\Theta_{\mathsf{m}}, (\mathcal{M}_w)')$$

taking $\mu = \mathsf{m}$ and $v_3 = \mathcal{M}_w$ in the local solver $(\mu) \Longrightarrow$

$$-(\Theta_{\mathsf{m}}, (\mathcal{M}_w)') = -\langle 1, (\widehat{\Theta}_{\mathsf{m}})(\mathcal{M}_w)n\rangle - (\mathcal{M}_{\mathsf{m}}, \mathcal{M}_w)$$

Then $\Longrightarrow$

$$-(\mathcal{M}_w, (\Theta_{\mathsf{m}})') = -\langle 1, (\mathcal{M}_w)(\Theta_{\mathsf{m}})n\rangle + \langle 1, (\widehat{\Theta}_{\mathsf{m}})(\mathcal{M}_w)n\rangle - (\mathcal{M}_{\mathsf{m}}, \mathcal{M}_w)$$

Thus we have $\Longrightarrow$

$$\begin{aligned} d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) + (\mathcal{M}_{\mathsf{m}}, \mathcal{M}_w) = &- \langle 1, (\mathcal{W}_{\mathsf{m}})(\mathcal{T}_w)n\rangle \qquad + (\mathcal{T}_w, (\mathcal{W}_{\mathsf{m}})') \\ &+ \langle \widehat{\Theta}_{\mathsf{m}} - \Theta_{\mathsf{m}}, (\mathcal{M}_w)n\rangle \quad + (\kappa\mathcal{U}_{\mathsf{m}}, \mathcal{T}_w) \end{aligned}$$

using the local solver (w) with $v_6 = \mathcal{W}_{\mathsf{m}} \Longrightarrow$

$$(\mathcal{T}_w, (\mathcal{W}_{\mathsf{m}})') = \langle 1, (\widehat{\mathcal{T}}_w)(\mathcal{W}_{\mathsf{m}})n\rangle + (\kappa\mathcal{N}_w, \mathcal{W}_{\mathsf{m}})$$

Then we have

$$\begin{aligned} d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) + (\mathcal{M}_{\mathsf{m}}, \mathcal{M}_w) = &\langle \widehat{\mathcal{T}}_w - \mathcal{T}_w, (\mathcal{W}_{\mathsf{m}})n\rangle \qquad + (\kappa\mathcal{N}_w, \mathcal{W}_{\mathsf{m}}) \\ &+ \langle \widehat{\Theta}_{\mathsf{m}} - \Theta_{\mathsf{m}}, (\mathcal{M}_w)n\rangle \quad + (\kappa\mathcal{U}_{\mathsf{m}}, \mathcal{T}_w) \end{aligned}$$

using the local solver (μ) with $\mu = \mathsf{m}$ and $v_2 = \mathcal{N}_w \Longrightarrow$

$$(\kappa \mathcal{W}_{\mathsf{m}}, \mathcal{N}_w) = -(\mathcal{U}_{\mathsf{m}}, (\mathcal{N}_w)') - d^2(\mathcal{N}_{\mathsf{m}}, \mathcal{N}_w)$$

IBP $\Longrightarrow$

$$-(\mathcal{U}_{\mathsf{m}}, (\mathcal{N}_w)') = -\langle 1, (\mathcal{U}_{\mathsf{m}})(\mathcal{N}_w)n\rangle + (\mathcal{N}_w, (\mathcal{U}_{\mathsf{m}})')$$

using the local solver (w) with $v_5 = \mathcal{U}_{\mathsf{m}} \Longrightarrow$

$$(\mathcal{N}_w, (\mathcal{U}_{\mathsf{m}})') = \langle 1, (\widehat{\mathcal{N}}_w)(\mathcal{U}_{\mathsf{m}})n\rangle - (\kappa \mathcal{T}_w, \mathcal{U}_{\mathsf{m}})$$

Then we have $\Longrightarrow$

$$-(\mathcal{U}_{\mathsf{m}}, (\mathcal{N}_w)') = \langle \widehat{\mathcal{N}}_w - \mathcal{N}_w, (\mathcal{U}_{\mathsf{m}})n\rangle - (\kappa \mathcal{T}_w, \mathcal{U}_{\mathsf{m}})$$

Thus $\Longrightarrow$

$$(\kappa \mathcal{U}_{\mathsf{m}}, \mathcal{N}_w) = \langle \widehat{\mathcal{N}}_w - \mathcal{N}_w, (\mathcal{U}_{\mathsf{m}})n\rangle - (\kappa \mathcal{T}_w, \mathcal{U}_{\mathsf{m}}) - d^2(\mathcal{N}_w, \mathcal{N}_{\mathsf{m}})$$

and hence,

$$(\kappa \mathcal{N}_w, \mathcal{W}_{\mathsf{m}}) + (\kappa \mathcal{U}_{\mathsf{m}}, \mathcal{T}_w) = \langle \widehat{\mathcal{N}}_w - \mathcal{N}_w, (\mathcal{U}_{\mathsf{m}})n\rangle - d^2(\mathcal{N}_w, \mathcal{N}_{\mathsf{m}})$$

Then $\Longrightarrow$

$$\begin{aligned} d^2(\mathcal{T}_w, \mathcal{T}_{\mathsf{m}}) + d^2(\mathcal{N}_w, \mathcal{N}_{\mathsf{m}}) + (\mathcal{M}_{\mathsf{m}}, \mathcal{M}_w) =& \langle \widehat{\Theta}_{\mathsf{m}} - \Theta_{\mathsf{m}}, (\mathcal{M}_w)n\rangle \\ &+ \langle \widehat{\mathcal{N}}_w - \mathcal{N}_w, (\mathcal{U}_{\mathsf{m}})n\rangle \\ &+ \langle \widehat{\mathcal{T}}_w - \mathcal{T}_w, (\mathcal{W}_{\mathsf{m}})n\rangle \end{aligned}$$

Thus $\Longrightarrow$

$$\begin{aligned}\langle\widehat{\Theta}_w, \mathsf{m}n\rangle = -\langle w, \mathcal{T}_{\mathsf{m}}n\rangle &+ \langle\widehat{\Theta}_{\mathsf{m}} - \Theta_{\mathsf{m}}, (\mathcal{M}_w)n\rangle \\ &+ \langle\widehat{\mathcal{N}}_w - \mathcal{N}_w, (\mathcal{U}_{\mathsf{m}})n\rangle \\ &+ \langle\widehat{\mathcal{T}}_w - \mathcal{T}_w, (\mathcal{W}_{\mathsf{m}})n\rangle \\ &+ \langle\widehat{\Theta}_w - \Theta_w, (\mathsf{m} - \mathcal{M}_{\mathsf{m}})n\rangle \\ &- \langle\widehat{\mathcal{N}}_{\mathsf{m}} - \mathcal{N}_{\mathsf{m}}, (\mathcal{U}_w)n\rangle \\ &- \langle\widehat{\mathcal{T}}_{\mathsf{m}} - \mathcal{T}_{\mathsf{m}}, (\mathcal{W}_w)n\rangle\end{aligned}$$

We do the same procedure to (B.2c) and get

$$\begin{aligned}\langle\widehat{\Theta}_{\mathsf{u}}, \mathsf{m}n\rangle = &d^2(\mathcal{T}_{\mathsf{u}}, \mathcal{T}_{\mathsf{m}}) \quad + d^2(\mathcal{N}_{\mathsf{u}}, \mathcal{N}_{\mathsf{m}}) + (\mathcal{M}_{\mathsf{u}}, \mathcal{M}_{\mathsf{m}}) \\ -\langle\mathsf{u}, \mathcal{N}_{\mathsf{m}}n\rangle \quad &+ \langle\widehat{\Theta}_{\mathsf{u}} - \Theta_{\mathsf{u}}, (\mathsf{m} - \mathcal{M}_{\mathsf{m}})n\rangle \\ &- \langle\widehat{\mathcal{N}}_{\mathsf{m}} - \mathcal{N}_{\mathsf{m}}, \mathcal{U}_{\mathsf{u}}n\rangle \\ &- \langle\widehat{\mathcal{T}}_{\mathsf{m}} - \mathcal{T}_{\mathsf{m}}, \mathcal{W}_{\mathsf{u}}n\rangle\end{aligned}$$

Similarly, from (B.2d) we get

$$\begin{aligned}d^2(\mathcal{T}_{\mathsf{m}}, \mathcal{T}_{\mathsf{u}}) + d^2(\mathcal{N}_{\mathsf{m}}, \mathcal{N}_{\mathsf{u}}) + (\mathcal{M}_{\mathsf{m}}, \mathcal{M}_{\mathsf{u}}) = &\langle\widehat{\Theta}_{\mathsf{m}} - \Theta_{\mathsf{m}}, (\mathcal{M}_{\mathsf{u}})n\rangle \\ &+ \langle\widehat{\mathcal{N}}_{\mathsf{u}} - \mathcal{N}_{\mathsf{u}}, (\mathcal{U}_{\mathsf{m}})n\rangle \\ &+ \langle\widehat{\mathcal{T}}_{\mathsf{u}} - \mathcal{T}_{\mathsf{u}}, (\mathcal{W}_{\mathsf{m}})n\rangle\end{aligned}$$

and then we get

$$\begin{aligned}\langle\widehat{\Theta}_{\mathsf{u}},\mathsf{m}n\rangle = -\langle\mathsf{u},\mathcal{N}_{\mathsf{m}}n\rangle &+ \langle\widehat{\Theta}_{\mathsf{m}}-\Theta_{\mathsf{m}},(\mathcal{M}_{\mathsf{u}})n\rangle\\ &+ \langle\widehat{\mathcal{N}}_{\mathsf{u}}-\mathcal{N}_{\mathsf{u}},(\mathcal{U}_{\mathsf{m}})n\rangle\\ &+ \langle\widehat{\mathcal{T}}_{\mathsf{u}}-\mathcal{T}_{\mathsf{u}},(\mathcal{W}_{\mathsf{m}})n\rangle\\ &+ \langle\widehat{\Theta}_{\mathsf{u}}-\Theta_{\mathsf{u}},(\mathsf{m}-\mathcal{M}_{\mathsf{m}})n\rangle\\ &- \langle\widehat{\mathcal{N}}_{\mathsf{m}}-\mathcal{N}_{\mathsf{m}},(\mathcal{U}_{\mathsf{u}})n\rangle\\ &- \langle\widehat{\mathcal{T}}_{\mathsf{m}}-\mathcal{T}_{\mathsf{m}},(\mathcal{W}_{\mathsf{u}})n\rangle\end{aligned}$$

Similarly, we can evaluate this identity to obtain an expression for (B.2e) and we get

$$\begin{aligned}\langle\widehat{\Theta}_{\mu},\mathsf{m}n\rangle = -\langle\mu,\Theta_{\mathsf{m}}n\rangle &+ \langle\widehat{\Theta}_{\mathsf{m}}-\Theta_{\mathsf{m}},(\mathcal{M}_{\mu})n\rangle\\ &+ \langle\widehat{\mathcal{N}}_{\mu}-\mathcal{N}_{\mu},(\mathcal{U}_{\mathsf{m}})n\rangle\\ &+ \langle\widehat{\mathcal{T}}_{\mu}-\mathcal{T}_{\mu},(\mathcal{W}_{\mathsf{m}})n\rangle\\ &+ \langle\widehat{\Theta}_{\mu}-\Theta_{\mu},(\mathsf{m}-\mathcal{M}_{\mathsf{m}})n\rangle\\ &- \langle\widehat{\mathcal{N}}_{\mathsf{m}}-\mathcal{N}_{\mathsf{m}},(\mathcal{U}_{\mathsf{u}})n\rangle\\ &- \langle\widehat{\mathcal{T}}_{\mathsf{m}}-\mathcal{T}_{\mathsf{m}},(\mathcal{W}_{\mathsf{u}})n\rangle\end{aligned}$$

From (B.2f) we get

$$\begin{aligned}\langle\widehat{\Theta}_{p},\mathsf{m}n\rangle = -(p,\mathcal{U}_{\mathsf{m}}) &+ \langle\widehat{\Theta}_{\mathsf{m}}-\Theta_{\mathsf{m}},(\mathcal{M}_{p})n\rangle\\ &+ \langle\widehat{\mathcal{N}}_{p}-\mathcal{N}_{p},(\mathcal{U}_{\mathsf{m}})n\rangle\\ &+ \langle\widehat{\mathcal{T}}_{p}-\mathcal{T}_{p},(\mathcal{W}_{\mathsf{m}})n\rangle\\ &+ \langle\widehat{\Theta}_{p}-\Theta_{p},(\mathsf{m}-\mathcal{M}_{\mathsf{m}})n\rangle\\ &- \langle\widehat{\mathcal{N}}_{\mathsf{m}}-\mathcal{N}_{\mathsf{m}},(\mathcal{U}_{p})n\rangle\\ &- \langle\widehat{\mathcal{T}}_{\mathsf{m}}-\mathcal{T}_{\mathsf{m}},(\mathcal{W}_{p})n\rangle\end{aligned}$$

Similarly, From (B.2g) we get

$$\begin{aligned}\langle\widehat{\Theta}_q, \mathsf{m}n\rangle = -(q, \mathcal{W}_{\mathsf{m}}) &+ \langle\widehat{\Theta}_{\mathsf{m}} - \Theta_{\mathsf{m}}, (\mathcal{M}_q)n\rangle \\ &+ \langle\widehat{\mathcal{N}}_q - \mathcal{N}_q, (\mathcal{U}_{\mathsf{m}})n\rangle \\ &+ \langle\widehat{\mathcal{T}}_q - \mathcal{T}_q, (\mathcal{W}_{\mathsf{m}})n\rangle \\ &+ \langle\widehat{\Theta}_q - \Theta_q, (\mathsf{m} - \mathcal{M}_{\mathsf{m}})n\rangle \\ &- \langle\widehat{\mathcal{N}}_{\mathsf{m}} - \mathcal{N}_{\mathsf{m}}, (\mathcal{U}_q)n\rangle \\ &- \langle\widehat{\mathcal{T}}_{\mathsf{m}} - \mathcal{T}_{\mathsf{m}}, (\mathcal{W}_q)n\rangle\end{aligned}$$

We can get similar results for (B.3a) – (B.3g) and (B.4a) – (B.4g).

REFERENCES

[1] D. N. Arnold, *Discretization by finite elements of a model parameter dependent problem*, Numer. Math. **37** (1981), 405–421.

[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal. **39** (2002), 1749–1779.

[3] D. N. Arnold, F. Brezzi, R. Falk, and L. D. Marini, *Locking-free Reissner–Mindlin elements without reduced integration*, Comput. Methods Appl. Mech. Engrg. **196** (2007), 3660–3671.

[4] D. N. Arnold, F. Brezzi, and D. Marini, *A family of discontinuous Galerkin finite elements for the Reissner–Mindlin plate*, J. Sci. Comput. **22** (2005), 25–45.

[5] H. Stolarski and T. Belytschko. *Membrane locking and reduced integration for curved elements*, J. Appl. Mech. **49** (1982), 172–176.

[6] D. N. Arnold and R. Falk, *A uniformly accurate finite element method for the Reissner–Mindlin plate*, SIAM J. Numer. Anal. **26** (1989), no. 6, 1276–1290.

[7] D. N. Arnold and F. Brezzi. *Locking-free finite element methods for shells*, Math. Comp. **66** (1997), 1–14.

[8] D. N. Arnold and R. S. Falk, *The boundary layer for the Reissner–Mindlin plate model*, SIAM J. Math. Anal. **21** (1990), no. 2, 281–312.

[9] D. N. Arnold and R. S. Falk, *Analysis of a linear-linear finite element for the Reissner–Mindlin plate model*, **7** (1997), no. 2, 217–238.

[10] D. N. Arnold and X. Liu, *Interior estimates for a low order finite element method for the Reissner–Mindlin plate model*, Adv. Comput. Math. **7** (1997), no. 3, 337–360.

[11] I. Babuška and M. Suri, *Locking effects in the finite element approximation of elasticity problems*, Numer. Math. **62** (1992), 439–463.

[12] I. Babuška and M. Suri, *On locking and robustness in the finite element method*, SIAM J. Numer. Anal. **29** (1992), 1261–1293.

[13] F. Brezzi and M. Fortin, *Numerical approximation of Mindlin-Reissner plates*, Math. Comp. **47**, no. 175.

[14] P. Castillo, B. Cockburn, D. Schötzau, and C. Schwab, *Optimal a priori error estimates for the hp-version of the local discontinuous Galerkin method for convection-diffusion problems*, Math. Comp. **71** (2002), 455–478.

[15] F. Celiker, *Discontinuous Galerkin methods for structural mechanics*, Ph.D. thesis, University of Minnesota, Minneapolis, 2005.

[16] F. Celiker and B. Cockburn, *Element-by-element post-processing of discontinuous Galerkin methods for Timoshenko beams*, J. Sci. Comput. **27** (2006), no. 1–3, 177–187.

[17] F. Celiker, B. Cockburn, S. Güzey, R. Kanapady, S.-C. Soon, H. K. Stolarski, and K. K. Tamma, *Discontinuous Galerkin methods for Timoshenko beams*, Numerical Mathematics and Advanced Applications, ENUMATH 2003, Springer, 2003, pp. 221–231.

[18] F. Celiker, B. Cockburn, and K. Shi, *A projection-based error analysis of HDG methods for Timoshenko beams, submitted*, Math. Comp., to appear.

[19] F. Celiker, B. Cockburn, and K. Shi, *Hybridizable discontinuous Galerkin methods for Timoshenko beams*, J. Sci. Comput. **44** (2010), 1–37.

[20] F. Celiker, B. Cockburn, and H.K. Stolarski, *Locking-free optimal discontinuous Galerkin methods for Timoshenko beams*, SIAM J. Numer. Anal. **44** (2006), no. 6, 2297–2325.

[21] C. Chinosi, C. Lovadina, and L.D. Marini, *Nonconforming locking-free finite elements for Reissner–Mindlin plates*, Comput. Methods Appl. Mech. Engrg. **195** (2006), 3448–3460.

[22] P. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[23] H. Stolarski and T. Belytschko. *Shear and membrane locking in curved C^0 elements*, Comput. Methods Appl. Mech. Engrg. **41** (1983) 279–296.

[24] F. Celiker and B. Cockburn. *Element-by-element post-processing of discontinuous Galerkin methods for Timoshenko beams*, J. Sci. Comput.

[25] R. Durán, A. Ghioldi, and N. Wolanski, *A finite element method for the Mindlin-Reissner plate model*, SIAM J. Numer. Anal. **28** (1991), 1004–1014.

[26] R. Durán, E. Herhández, L. Hervella-Nieto, E. Liberman, and R. Rodríguez, *Error estimates for low-order quadrilateral finite elements for plates*, SIAM J. Numer. Anal. **41** (2003), 1751–1772.

[27] R.S. Falk and T. Tu, *Locking-free finite elements for the Reissner–Mindlin plate*, Math. Comp. **69** (2000), no. 231, 911–928.

[28] P. Houston, C. Schwab, and E. Süli, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal. **39** (2002), 2133–2163.

[29] T. J. R. Hughes, R. L. Taylor, and W. Kanoknukulchai, *A simple and efficient element for plate bending*, Internat. J. Numer. Methods Engrg. **11** (1977), 1529–1543.

[30] L. Li, *Discretization of the Timoshenko beam problem by the p and the h-p versions of the finite element method*, Numer. Math. **57** (1990), 413–420.

[31] C. Lovadina, *A low-order nonconforming finite element for Reissner–Mindlin plates*, SIAM J. Numer. Anal. **42** (2005), no. 6, 2688–2701.

[32] D. S. Malkus and T. J. R Hughes, *Mixed finite element methods-reduced integration and selective integration techniques: a unification of concepts*, Comput. Methods Appl. Mech. Engrg. **15** (1978), 63–81.

[33] J. Pitkäranta and M. Suri, *Upper and lower bounds for plate-bending finite elements*, Numer. Math. **84** (2000), 611–648.

[34] C. Schwab, *p- and hp-FEM. Theory and application to solid and fluid mechanics*, Oxford University Press, New York, 1998.

[35] M. Suri, I. Babuška, and C. Schwab, *Locking effects in the finite element approximation of plate models*, Math. Comp. **210** (1995), 461–482.

[36] S. Zhang, *An asymptotic analysis on the form of Naghdi type arch model*, Math. Models Methods Appl. Sci. **18** (2008), no. 3.

[37] Z. Zhang, *Arch beam models: finite element analysis and superconvergence*, Numer. Math. **61** (1992), 117–143.

[38] Z. Zhang, *A note on the hybrid-mixed C^0 curved beam elements*, Comput. Methods Appl. Mech. Engrg. **95** (1992), 243–252.

[39] Z. Zhang, *Locking and robustness in the finite element method for circular arch problem*, Numer. Math. **69** (1995), 509–522.

[40] Z. Zhang and S. Zhang, *Wilson's element for the Reissner–Mindlin plate*, Comput. Methods Appl. Mech. Engrg. **113** (1994), 55–65.

[41] Z. Zhang and S. Zhang, *Derivative superconvergence of rectangular finite elements for the Reissner–Mindlin plate*, Comput. Methods Appl. Mech. Engrg. **134** (1996), 1–16.

[42] F. Celiker, L. Fan, S. Zhang, and Z. Zhang, *Locking-free optimal discontinuous Galerkin methods for a Naghdi-type arch model*, J. Sci. Comput. **52** (2012), 49–84.

[43] B. Cockburn, J. Gopalakrishnan, and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal. **47** (2009), no. 2, 1319–1365.

[44] B. Cockburn, B. Dong, and J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp. **77** (2008), no. 264, 1887–1916.

[45] B. Cockburn, J. Guzmán, and H. Wang, *Superconvergent discontinuous Galerkin methods for second-order elliptic problems*, Math. Comp. **78** (2009), 1–24.

[46] F. Celiker,B. Cockburn, *A projection-based error analysis of HDG methods for Timoshenko beams*, Math. Comp. **81** (2012), 131–151.

[47] B. Cockburn, J. Gopalakrishnan, and F.-J. Sayas, *A projection-based error analysis of HDG methods*, Math. Comp. **79** (2010), no. 271, 1351–1367.

[48] B. Cockburn, J. Gopalakrishnan, N. C. Nguyen, J. Peraire, and F.-J. Sayas, *Analysis of HDG methods for Stokes flow*, Math. Comp. **80** (2011), no. 274, 723–760.

[49] B. Cockburn, B. Dong, J. Guzmán, M. Restelli, and R. Sacco, *A hybridizable discontinuous Galerkin method for steady-state convection-diffusion problems*, SIAM J. Sci. Comput. **31** (2009), no. 5, 3827–3846.

[50] F. Celiker, L. Fan, and Z. Zhang, *Element-by-element post-processing of DG methods for Naghdi arches.*, Int. J. Numer. Anal. Model. **8** (2011), no. 3, 391–409.

[51] F. Celiker,B. Cockburn, *Superconvergence of the numerical traces of discontinuous Galerkin and hybridized methods for convection-diffusion problems in one space dimension*, Math. Comp. **76** (2007), no. 257, 67–96.

[52] B. Cockburn and R. Ichikawa, *Adjoint recovery of superconvergent linear functionals from Galerkin approximations*, J. Sci. Comput. **32** (2007), no. 2, 201–232.

[53] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO, Anal. Numér. **19** (1985), 611–643.

[54] V. Thomée, *Galerkin Finite Element Methods for parabolic equations*, Springer Verlag, 1997.

[55] W.H. Reed and T.R. Hill, *Triangular mesh methods for the neutron transport equation*, Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[56] P. Lesaint and P. A. Raviart, *On a finite element method for solving the neutron transport equation*, Mathematical aspects of finite elements in partial differential equations (C. de Boor, ed.), Academic Press, 1974, pp. 89–145.

[57] M. Delfour, W. Hager, and F. Trochu, *Discontinuous Galerkin methods for ordinary differential equations*, Math. Comp. **36** (1981), 455–473.

[58] B. Cockburn and B. Dong, *An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems*, J. Sci. Comput. **32** (2007), no. 2, 233–262.

[59] B. Cockburn, G. Kanschat, I. Perugia, and D. Schötzau, *Superconvergence of the local discontinuous Galerkin method for elliptic problems on Cartesian grids*, SIAM J. Numer. Anal. **39** (2001), 264–285.

[60] B. Dong and C-W. Shu, *Analysis of a local discontinuous Galerkin method for linear time-dependent fourth-order problems*, SIAM J. Numer. Anal. **47** (2009), no. 5, 3240–3268.

[61] D. Schötzau and C. Schwab, *Time discretization of parabolic problems by the hp-version of the Discontinuous Galerkin Finite Element Method*, SIAM J. Numer. Anal. **38** (2000), 837–875.

[62] H. Parish. *A critical survey of the 9-node degenerated shell element with special emphasis on this shell application and reduced integration*, Comput. Methods Appl. Mech. Engrg. **20** (1979), 323–350.

[63] Z. Xie and Z. Zhang, *Uniform superconvergence analysis of the discontinuous Galerkin method for a singularly perturbed problem in 1-D*, Math. Comp. **79** (2010), no. 269, 35–45.

[64] K. Bey and J. T. Oden, *hp-version discontinuous Galerkin methods for hyperbolic conservation laws*, Comput. Methods Appl. Mech. Engrg. **133** (1996), 259–286.

[65] K. Bey, J. T. Oden, and A. Patra, *A parallel hp-adaptive discontinuous Galerkin method for hyperbolic conservation laws*, Appl. Numer. Math. **20** (1996), 321–286.

[66] K. Bey, A. Patra, and J. T. Oden. *hp-version discontinuous Galerkin methods for hyperbolic conservation laws: A parallel strategy*, Internat. J. Numer. Methods Engrg. **38** (1995), 3889–3908.

[67] R. Biswas, K. Devine, and J. Flaherty. *Parallel, adaptive finite element methods for conservation laws*, Appl. Numer. Math. **14** (1994), 255–283.

[68] B. Cockburn, S. Lin, and C.-W. Shu. *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: One dimensional systems*, J. Comput. Phys. **84** (1989), 90–113.

[69] B. Cockburn and C.-W. Shu. *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws II: General framework*, Math. Comp. **52** (1989), 411–435.

[70] B. Cockburn and C.-W. Shu. *The Runge-Kutta local projection P^1-discontinuous Galerkin method for scalar conservation laws*, RAIRO Modél. Math. Anal. Numér. **25** (1991), 337–361.

[71] K. Devine and J. Flaherty. *Parallel adaptive hp-refinement techniques for conservation laws*, Appl. Numer. Math. **20** (1996), 367–386.

[72] K. Eriksson and C. Johnson. *Adaptive finite element methods for parabolic problems i: a linear model problem*, SIAM J. Numer. Anal. **28** (1991), 12–23.

[73] D. G. Ashwell and A. B. Sabir. *Limitations of certain curved finite elements when applied to arches*, Internat. J. Mech. Sci. **13** (1971), 133–139.

[74] S. W. Lee and T. H. H. Pian. *Improvements of plate and shell finite elements by mixed formulations*, AIAA Journal. **16** (1978), 29–34.

[75] K. Eriksson and C. Johnson. *Adaptive finite element methods for parabolic problems ii: optimal error estimates in $l_\infty l_2$ and $l_\infty l_\infty$*, SIAM J. Numer. Anal. **32** (1995), 706–740.

[76] K. Eriksson, C. Johnson, and V. T. ee. *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO, Anal. Numér. **19** (1985), 611–643.

[77] F. Bassi and S. Rebay. *A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations*, J. Comput. Phys. **131** (1997), 267–279.

[78] H. Stolarski, T. Belytschko, and S.-H. Lee. *A review of shell finite elements and corotational theories*, Comput. Mech. Advances. **2** (1995), 125–212.

[79] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini. *A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows*, 2nd European Conference on Turbomachinery Fluid Dynamics and Thermodynamics, 99–108, Antwerpen, Belgium, March 5–7 1997.

[80] C. E. Baumann and J. T. Oden. *A discontinuous hp-finite element method for the Navier-Stokes equations*, 10th. Intern. Confer. on Finite Element in Fluids, 1998.

[81] C. E. Baumann and J. T. Oden. *A discontinuous hp-finite element method for convection-diffusion problems*, Comput. Methods Appl. Mech. Engrg. **175** (1999), 311–341.

[82] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. *An a priori error analysis of the local discontinuous Galerkin method for elliptic problems*, SIAM J. Numer. Anal. **38** (2000), 1676–1706.

[83] B. Cockburn, G. Kanschat, I. Perugia, and D. Schötzau. *Local discontinuous Galerkin methods for elliptic problems*, Commun. Numer. Meth. Engng. **18** (2002), 69–75.

[84] B. Cockburn, G. Kanschat, and D. Schötzau. *Local discontinuous Galerkin methods for the Oseen equations*, Math. Comp. **73** (2004), 569–593.

[85] B. Cockburn, G. Kanschat, D. Schötzau, and C. Schwab. *Local discontinuous Galerkin methods for the Stokes system*, SIAM J. Numer. Anal. **40** (2002), 319–343.

[86] Sheng. Zhang, *A Linear Shell Theory Based on Variational Principles*, Ph.D. thesis, Pennsylvania State University, 2001.

[87] B. Rivière, M. F. Wheeler, and V. Girault. *A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems*, SIAM J. Numer. Anal. **39** (2001), 902–931.

[88] M. F. Wheeler. *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal. **15** (1978), 152–161.

[89] B. Cockburn, G. Karniadakis, and C.-W. Shu. *The development of discontinuous Galerkin methods*. Lect. Notes Comput. Sci. Engrg., pages 3–50, Berlin, February 2000. Springer Verlag.

[90] O. Zienkiewicz, R. L. Taylor, and J. M. Too. *Reduced integration technique in general analysis of plates and shells*. Internat. J. Numer. Methods Engrg. **3** (1971), 275–290.

ABSTRACT

DG AND HDG METHODS FOR CURVED STRUCTURES

by

LI FAN

December 2013

Advisor: Dr. Fatih Celiker

Co-Advisor: Dr. Zhimin Zhang

Major: Mathematics

Degree: Doctor of Philosophy

We introduce and analyze discontinuous Galerkin methods for a Naghdi type arch model. We prove that, when the numerical traces are properly chosen, the methods display optimal convergence uniformly with respect to the thickness of the arch. These methods are thus free from membrane and shear locking. We also prove that, when polynomials of degree k are used, *all* the numerical traces superconverge with a rate of order h^{2k+1}.

Based on the superconvergent phenomenon and we show how to post-process them in an element-by-element fashion to obtain a far better approximation. Indeed, we prove that, if polynomials of degree k are used, the post-processed approximation converges with order $2k+1$ in the L^2-norm throughout the domain. This has to be contrasted with the fact that before post-processing, the approximation converges with order $k+1$ only. Moreover, we show that this superconvergence property does not deteriorate as the thickness of the arch becomes extremely small.

Since the DG methods suffer from too many degree of freedoms we introduce and analyze a class of hybridizable discontinuous Galerkin (HDG) methods for Naghdi arches. The main

feature of these methods is that they can be implemented in an efficient way through a hybridization procedure which reduces the globally coupled unknowns to approximations to the transverse and tangential displacement and bending moment at the element boundaries. The error analysis of the methods is based on the use of a projection especially designed to fit the structure of the numerical traces of the method. This property allows to prove in a very concise manner that the projection of the errors is bounded in terms of the distance between the exact solution and its projection. The study of the influence of the stabilization function on the approximation is then reduced to the study of how they affect the approximation properties of the projection in a single element. Consequently, we prove that HDG methods have the same result as DG methods.

At the end of the thesis, we talk a little bit of shell problems.

AUTOBIOGRAPHICAL STATEMENT

LI FAN

Education

- 2013, Ph.D. in Mathematics (*expected*)
  Wayne State University, Detroit, Michigan, USA.
- 2008, M.S. in Mathematics
  University of Science and Technology of China, Hefei, Anhui, China.
- 2004, B.S. in Mathematics
  University of Science and Technology of China, Hefei, Anhui, China.

Papers and preprints

1. F. Celiker, L. Fan, S. Zhang, and Z. Zhang, *Locking-free optimal discontinuous Galerkin methods for a Naghdi-type arch model*. Journal of Scientific Computing, July 2012, Volume 52, pp 49-84.
2. F. Celiker, L. Fan, and Z. Zhang, *Element-by-element post-processing of discontinuous Galerkin methods for Naghdi arches*. Int. J. of Numer. Anal. and Model, 8(2011) 391-409.
3. F. Celiker, L. Fan, *Hybridizable discontinuous Galerkin methods for Naghdi arches*. Journal of Scientific Computing, published online.
4. Huiqing. Zhu, F. Celiker, L. Fan, *Element-by-element post-processing of discontinuous Galerkin methods for Singularly Perturbed Problems*. In preparation.